

**ÉTUDE DES RIBOREGULATEURS GLYCINE CHEZ**  
***CLOSTRIDIUM DIFFICILE***

Par

Auréliane Michaud  
Programme de microbiologie

Mémoire présenté à la Faculté de médecine et des sciences de la santé en vue de l'obtention  
du grade de maître ès sciences (M.Sc.) en microbiologie

Sherbrooke, Québec, Canada

Mai 2017

Membres du jury d'évaluation

Louis-Charles Fortier, département de microbiologie et infectiologie, FMSS  
Daniel Lafontaine, département de biologie, faculté des sciences  
Alfredo Menendez, département de microbiologie et infectiologie, FMSS  
Vincent Burrus, département de biologie, faculté des sciences

© Michaud Auréliane, 2017

# Étude des riborégulateurs glycine chez *Clostridium difficile*

Par

Auréliane Michaud  
Programmes de microbiologie

Mémoire présenté à la Faculté de médecine et des sciences de la santé en vue de l'obtention du diplôme de maître ès sciences (M.Sc.) en microbiologie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4

*Clostridium difficile* est une bactérie pathogène opportuniste causant chaque année des ravages dans les hôpitaux. Depuis quelques années, elle est devenue de plus en plus résistante aux antibiotiques et représente un coût financier grandissant pour le système de santé qui peine à la traiter. De nouvelles stratégies sont donc nécessaires afin de lutter contre *C. difficile*. Une solution potentielle serait l'utilisation des riborégulateurs comme cible antibiotique étant donné que ces structures d'ARN permettent la régulation de gènes bactériens suite à la liaison d'un ligand spécifique. En ciblant ces structures par l'utilisation d'un analogue non métabolisable, il pourrait être possible d'impacter significativement la viabilité ou la compétitivité d'une bactérie. Durant cette maîtrise, l'étude s'est portée sur les deux riborégulateurs liant la glycine, un acide aminé très présent dans le tractus digestif et important dans le métabolisme bactérien.

Ces riborégulateurs (RR), RR-symp contrôlant un symporteur sodium alanine ou sodium/glycine et RR-grdX contrôlant une glycine réductase n'avaient été identifiés que par bio-informatique. Une première approche structurale par extension d'amorce a été nécessaire afin de déterminer les structures primaires et secondaires des riborégulateurs. Il a ainsi pu être confirmé que le RR-symp était un riborégulateur en tandem comportant deux aptamères suivis d'une seule plateforme d'expression tandis que le RR-grdX est un singlet de type 2, doté d'un seul aptamère mais également d'une longue tige-boucle en amont. Des analyses avec des essais rapporteurs ( $\beta$ -galactosidase avec des constructions chez *B. subtilis*) ont permis de constater qu'en présence de glycine le RR-symp était un répresseur alors que le RR-grdX était quant à lui un activateur. Des mutants de liaison ont permis d'abolir toute régulation et de confirmer la spécificité de liaisons aux RR. Des tests pour déterminer l'affinité des RR pour la glycine *in vitro* ont confirmé ces résultats et ont également permis de montrer dans le cas de RR-grdX que la tige-boucle en amont était importante pour l'activité

Par contre l'inactivation des gènes régulés par les riborégulateurs n'a pas eu d'impact significatif sur la croissance de la bactérie que ce soit en milieu riche ou en milieu pauvre, ce qui semble indiquer que ces riborégulateurs ne semblent pas, à priori, être des cibles intéressantes dans une optique thérapeutique. Ces travaux ont néanmoins permis de caractériser davantage ces structures encore méconnues chez *C. difficile*.

**Mots clefs :** *Clostridium difficile*, riborégulateur, glycine, singlet, tandem.

## Table des matières

Résumé .....	ii
Liste des abréviations .....	vi
Liste des figures .....	ix
Liste des tableaux .....	x
Liste des annexes .....	x
I. Expression génétique – Introduction .....	1
I.1. Quelques notions historiques.....	1
I.2. L'expression génétique.....	2
II. Régulation génétique bactérienne .....	5
II.1. Régulation de la transcription .....	6
II.1.1. Les facteurs sigma .....	6
II.1.2. Réactions enzymatiques.....	7
II.1.3. Élongation.....	8
II.1.4. Terminaison.....	9
II.2. Régulation de la traduction .....	10
II.2.1. Initiation de la traduction .....	10
II.2.2. Modification de la stabilité des ARNm .....	11
II.2.3. Modification ou dissociation des ribosomes.....	12
II.3. Régulation métabolique (post-traductionnelle).....	12
II.4. Antibiotiques et résistances .....	13
III. Riborégulateurs .....	16
III.1. Historique .....	16
III.2. Anatomie et fonctionnement.....	17
III.2.1. Régulation transcriptionnelle.....	17
III.2.2. Régulation traductionnelle.....	18
III.2.3. Régulation par dégradation (ribozyme) .....	19
III.2.4. Régulation de l'épissage alternatif .....	19
III.2.5. Régulation anti-sens .....	21
III.3. Classification.....	21
III.4. Riborégulateurs glycine .....	25
III.4.1. Structure.....	26
III.4.2. Coopérativité .....	28
III.4.3. Liaison du ligand et repliement .....	29

III.4.4. Aptamères fantômes.....	30
IV. Clostridium difficile.....	31
IV.1. Organisme .....	31
IV.2. Problématiques de santé publique .....	32
IV.2.1. Historique et chiffres.....	32
IV.2.1. Infection et toxines .....	33
IV.2.2. Pathologies et traitements.....	36
IV.3. Clostridium difficile souche 630 et riborégulateurs .....	41
V. Objectifs de maîtrise.....	43
VI. Matériels & Méthodes.....	44
VI.1. Souches utilisées .....	44
VI.2. Préparation des cellules compétentes.....	44
VI.3. Extraction d'ADN génomique.....	45
VI.4. Réaction de polymérisation en chaine (PCR) .....	46
VI.5. Construction de plasmides : Gibson et ligation.....	47
VI.5.1 Gibson .....	47
VI.5.2. Ligation.....	49
VI.6. Extraction rapide de plasmide à la silice (Miniprep).....	49
VI.6.1. Chez E. coli.....	49
VI.6.2. Chez C. difficile .....	50
VI.7. ClosTron .....	51
VI.8. Transformation et conjugaison .....	52
VI.8.1. Transformation.....	52
VI.8.2. Conjugaison.....	53
VI.9. Courbes de croissance.....	54
VI.10. Essais $\beta$ -glucuronidase chez C. difficile .....	55
VI.10.1. Essais $\beta$ -glucuronidase sur extraits cellulaires .....	56
VI.10.2. Essais $\beta$ -glucuronidase sur cellules perméabilisées .....	57
VI.11. Essais $\beta$ -galactosidase chez B. subtilis.....	57
VI.12. Extraction d'ARN de C. difficile .....	59
VI.13. Extension d'amorce (Primer extension).....	60
VI.14. Cartographie chimique (In-Line Probing) .....	62
VI.15. Transcription in vitro à cycle unique (Single Round in vitro transcription).....	64
VII. Résultats : Caractérisation du riborégulateur contrôlant le symporteur glycine/sodium chez C. difficile .....	66
VII.1. Contexte .....	66



VII.2. Résultats.....	68
VII.2.1. Le RR-symp, un riborégulateur avec une structure secondaire typique.....	68
VII.2.2. Une originalité fonctionnelle .....	70
VII.2.3. Une affinité en accord avec la littérature .....	74
VII.2.4. Un mutant à la croissance retardée.....	76
VIII. Résultats : Caractérisation du riborégulateur contrôlant la glycine réductase grdX chez C. difficile	78
VIII.1. Contexte.....	78
VIII.2. Résultats.....	80
VIII.2.1. RR-grdX, un riborégulateur structurellement inhabituel .....	80
VIII.2.2. Une fonctionnalité renouant avec la normale .....	82
VIII.2.3. Un singlet aussi efficace qu'un tandem .....	83
VIII.2.4. Le RR-grdX, un candidat à écarter.....	86
IX. Discussion.....	88
X. Conclusions.....	96
XI. Annexes .....	97
Remerciements .....	102
Liste des Références .....	103

## Liste des abréviations

A	Adénine
ADN	Acide désoxyribonucléique
AMPc	Adénosine monophosphate cyclique
ATP	Adénosine triphosphate
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messenger
ARNnc	Acide ribonucléique non codant
ARNr	Acide ribonucléique ribosomal
ARNt	Acide ribonucléique de transfert
BHI	Infusion cerveau-cœur ( <i>Brain heart infusion</i> )
BHIS	Infusion cerveau-cœur ( <i>Brain heart infusion supplemented</i> ) avec cystéine
CDC	<i>Centers for Disease Control</i>
°C	Degré Celsius
C	Cytosine
CAP	Protéine activatrice des gènes cataboliques ( <i>Catabolic activating protein</i> )
Cm	Chloramphénicol
CTD	Domaine C-Terminal ( <i>C-Terminal Domain</i> )
Ctrl	Contrôle
CPR	Complexe Protéique Ribonucléaire
DACD	Diarrhée à <i>Clostridium difficile</i>
dG	2'-déoxyguanosine
DNase	Désoxyribonucléase
DO	Densité optique
DTT	Dithiothréitol
EC	Complexe d'élongation ( <i>Elongation complex</i> )
EDTA	Acide éthylène-diamine-tétraacétique édétique ( <i>Ethylenediaminetetraacetic acid</i> )
Erm	Erythromycine

Fig.	Figure
FL	<i>Full Length</i>
g	Gra mMe
G	Guanine
GTP	Guanosine triphosphate
ICD	Infection à <i>Clostridium difficile</i>
INSPQ	Institut National de Santé Publique du Québec
Kd	Constante de dissociation
L	Ligand
LB	Luria-Bertani
M	Molaire
MM	Milieu minimum
mg	Milligra mMe
miARN	Micro acide ribonucléique
min	Minute
mL	Millilitre
MOPS	3-(N-morpholino)propanesulfonic acid
Mpb	Millions de paires de bases
nm	Nanomètre
nt	Nucléotide
NTD	Domaine N-Terminal ( <i>N-Terminal Domain</i> )
O/N	Durant la nuit ( <i>over night</i> )
OMS	Organisation Mondiale de la Santé
ONPG	Ortho-Nitrophényl- $\beta$ -galactoside
ORF	Cadre ouvert de lecture ( <i>Open Reading Frame</i> )
P1, P2, ...	Tige 1 ( <i>Paired 1</i> ), 2 ...
pARN	Petit acide ribonucléique
pb	Paire de bases
PBS	Tampon phosphate salin ( <i>Phosphate-buffered saline</i> )
PCI	Phénol Chloroforme Inositol

PCR	Réaction en chaîne par polymérase ( <i>Polymerase Chain Reaction</i> )
pH	Potentiel hydrogène
PNK	Polynucléotide Kinase
R	Résistant
RBS	Site de fixation du ribosome ( <i>Ribosome Binding Site</i> )
RNase	Ribonucléase
rNTP	Ribonucléotide triphosphate
rpm	Rotations par minute
S	Sensible
SD	Shine-Dalgarno
SDS	Sodium Dodécyl Sulfate
sec	Seconde
SHAPE	Selective 2'-hydroxyl acylation analyzed by primer extension
SM	<i>Starvation medium</i>
T	Thymine
T50	Concentration minimale de ligand pour avoir 50% de terminaison
Tab.	Tableau
TPP	Thiamine pyrophosphate
TY	<i>Tryptone yeast</i>
U	Uracile
μL	Microlitre
URSS	Union des Républiques Socialistes Soviétiques
UTR	Région non traduite ( <i>Untranslated region</i> )
VC	<i>Vibrio cholerae</i>
V <sub>f</sub>	Volume final
V <sub>max</sub>	Vitesse maximale
YIPP	Yeast inorganic pyrophosphatase
WT	Souche sauvage ( <i>Wild type</i> )
<sup>32</sup> P	Isotope radioactif du phosphore

## Liste des figures

Figure 1. Principe du dogme central de la biologie moléculaire. ....	2
Figure 2. Schématisation des possibilités principales de régulation de l'expression génétique chez la bactérie.....	5
Figure 3. Chronologie de la mise sur le marché de quelques antibiotiques (haut) et l'apparition de la première résistance à celui-ci (bas). Données CDC. ....	13
Figure 4. Cibles les plus communes des antibiotiques et stratégies de résistances. Adapté de (Wright, 2010).....	14
Figure 5. Schéma de la structure d'un ARNm bactérien doté d'un riborégulateur.....	16
Figure 6. Différents mécanismes utilisés pour la régulation de la transcription et de la traduction par les riborégulateurs (adapté de (Serganov and Nudler, 2013)). ....	18
Figure 7. Mécanisme de régulation par clivage de l'ARNm. Adapté de (Serganov and Nudler, 2013). ....	19
Figure 8. Mécanisme de régulation par épissage alternatif chez <i>C. difficile</i> . Adapté de (Serganov and Nudler, 2013).....	20
Figure 9. Exemples de riborégulateurs en tandem .....	27
Figure 10. Structure secondaire des aptamères du riborégulateur glycine chez <i>V. cholerae</i> (Huang et al., 2010). ....	27
Figure 11. Liaison coopérative des deux aptamères du riborégulateur glycine chez <i>V. cholerae</i> (VC). ....	28
Figure 12. Structures secondaires et tertiaires du riborégulateur glycine. Adapté de (Ruff and Strobel, 2014).....	29
Figure 13. Architecture des riborégulateurs glycine : en tandem (gauche), en singlet (type 1 centre, type 2 à droite). Adapté de (Ruff et al., 2016).....	30
Figure 14. Germination et production de toxines par <i>C. difficile</i> en présence d'un microbiote sain ou perturbé. (Adapté de (Taur and Pamer, 2014)).....	35
Figure 15. Carte montrant la dispersion des deux lignées épidémiques de <i>C. difficile</i> résistantes aux fluoroquinolones à travers le monde (He et al., 2013). ....	37
Figure 16. Principe de la réaction Gibson. Adapté de (Gibson et al., 2009).....	48
Figure 17. Principe du ClosTron. Adapté de (Kuehne and Minton, 2012).....	51
Figure 18. Réaction catalysée par la $\beta$ -D-glucuronidase. ....	55
Figure 19. Réaction catalysée par la $\beta$ -galactosidase. ....	58
Figure 20. Principe d'une cartographie chimique. ....	63
Figure 21. Principe de la transcription in vitro à cycle unique. ....	65
Figure 22. Structure et fonctionnement prédits du RR-symp (A) et données de RNAseq (B).....	66
Figure 23. Réaction de la voie de biosynthèse III de la glycine à partir d'alanine. ....	67
Figure 24. Extension d'amorce pour le RR-symp et séquence des aptamères (jaune) .....	69
Figure 25. Comparaison entre les structures primaire et secondaire du RR de <i>V. cholerae</i> et le RR-symp de <i>C. difficile</i> .....	70
Figure 26. Optimisation de la perméabilisation des cellules pour les essais rapporteurs .....	72
Figure 27. Activité glucuronidase du RR-symp chez <i>C. difficile</i> via perméabilisation .....	73
Figure 28. Activité galactosidase du RR-symp chez <i>B. subtilis</i> .....	74
Figure 29. Cartographie chimique de l'aptamère 2 du RR-symp et identification en couleur des zones où le clivage survient associé à une concentration croissante en ligand (haut), T50 sur la zone 1' (bas). ....	75
Figure 30. Courbes de croissance de la souche WT et le mutant inactivé du symp en milieu riche (A) et milieu pauvre (B). N=3.....	76

Figure 31. Fonctionnement prédit du riborégulateur grdX (A) et données de RNAseq (B), l'opéron potentiel de grdX est présenté en (C) où les autres gènes sont également impliqués dans la réductase en codant pour des sous-unités notamment. ....	78
Figure 32. Schéma simplifié de la réaction de Stickland .....	79
Figure 33. Extension d'amorce pour l'ARN du riborégulateur grdX. ....	80
Figure 34. Comparaison des structures primaire et secondaires entre le RR-symp et le RR-grdX de C. difficile. ....	81
Figure 35. Activité galactosidase du RR-grdX chez B. subtilis.....	82
Figure 36. Activité galactosidase du mutant U168A de liaison du RR-grdX.chez B. subtilis .....	83
Figure 37. Tests pour déterminer la meilleure concentration en rNTPs. ....	84
Figure 38. Gradient de concentration de glycine et détermination du T50 pour le RR-grdX.....	84
Figure 39. Transcription in vitro de la souche sauvage (WT) et du mutant de liaison (U168A) en absence (-L) ou en présence de glycine (+L).....	85
Figure 40. Transcription in vitro pour le mutant de repliement en fonction d'une concentration croissante de glycine. ....	85
Figure 41. Détermination du T50 pour le mutant de repliement du RR-grdX.....	86
Figure 42. Courbes de croissance de la souche sauvage (WT) et du mutant inactivé de grdX en milieu riche. N=3 .....	86
Figure 43. Courbes de croissance de la souche sauvage et du mutant inactivé de grdX en milieu pauvre. N=3 .....	87

## Liste des tableaux

Tableau 1. Bases puriques .....	3
Tableau 2. Bases pyrimidiques .....	3
Tableau 3. Exemples de facteurs sigma de la famille 70 .....	7
Tableau 4. Riborégulateurs liant des acides aminés. ....	22
Tableau 5. Riborégulateurs liant des bases azotées.....	22
Tableau 6. Riborégulateurs liant des coenzymes et des cofacteurs.....	23
Tableau 7. Riborégulateurs liant des ligands divers. ....	24
Tableau 8. Évolution du nombre de décès, de la létalité et des autres complications parmi les DACD nosocomiales (cat. 1a + 1b), Québec, 2011-2012 à 2015-2016 [N (%)] (données INSPQ 2016). ....	33
Tableau 9. Répartition des cas de DACD selon l'origine présumée de l'acquisition, Québec, 2015-2016 (N, %) (données INSPQ 2016) .....	34
Tableau 10. Traitement conseillé en fonction de la sévérité de l'infection .....	38
Tableau 11. Coût des antibiotiques en fonction de leur mode d'administration .....	38
Tableau 12. Conditions de PCR utilisées. X différant selon le Tm des couples d'amorces utilisés .....	47
Tableau 13. Amorces utilisées pour confirmer la présence de l'intron II et l'excision de l'intron I.....	54
Tableau 14. Amorces utilisées pour les extensions d'amorces .....	61
Tableau 15. Mélanges de séquençage correspondant aux différents nucléotides.....	61

## Liste des annexes

Annexe 1. Carte du plasmide pMTL007C-E5 pour CloStron.....	97
Annexe 2. Carte du plasmide pRPF144 .....	98
Annexe 3. Carte du plasmide pDG1661.....	99
Annexe 4. Structure secondaire du riborégulateur RR-symp annotée .....	100
Annexe 5. Structure secondaire du riborégulateur glycine RR-grdX annotée .....	101

## I. Expression génétique – Introduction

### I.1. Quelques notions historiques

Pendant des millénaires, la transmission des caractères au sein d'une même espèce a donné lieu à toutes les spéculations, qu'elles soient sensées ou complètement fantaisistes. Si Mendel a avancé les bases de l'hérédité en 1865 avec ses célèbres expériences sur les pois, il n'avait pas les clefs pour comprendre les mécanismes impliqués. Ses travaux tombèrent dans l'oubli avant d'être redécouverts en 1900 notamment par William Bateson, le premier à parler de génétique en 1905. Cette nouvelle science avait pourtant déjà quelques bases. En effet, l'acide désoxyribonucléique (ADN) fut isolé en 1869 par le biologiste suisse Friedrich Miescher à partir du pus de bandages chirurgicaux (il l'avait à l'origine appelée « nucléine » car retrouvée dans le noyau des cellules) (Dahm, 2008). L'identification des nucléotides (base, ose et groupement phosphate) remonte quant à elle à 1919 par le biologiste américain Phoebus Levene (Levene, 1919).

Il fallut toutefois attendre une vingtaine d'années pour que les avancées technologiques (dont la cristallographie aux rayons X montrant la structure ordonnée de l'ADN (Astbury and Bell, 1938)) permettent une meilleure observation et donc une meilleure compréhension de la génétique. C'est ainsi qu'en 1938, le terme de « biologie moléculaire » fit son apparition lorsque Warren Weaver l'utilisa pour décrire une nouvelle discipline de la science dans laquelle « de délicates nouvelles techniques sont utilisées pour enquêter sur encore plus de petits détails de certains processus de la vie » (Weaver, 1970). Cette définition un peu vague permit le regroupement de sciences auparavant bien distinctes telles la biochimie, la microbiologie, la physique et la génétique.

Toutefois ce n'est qu'en 1943 que l'ADN fut reconnu comme vecteur de l'information génétique (Avery et al., 1944). Puis, 10 ans plus tard, sa structure secondaire fut élucidée. En effet, se basant sur les travaux de Rosalind Franklin, Raymond Gosling et d'Erwin Chargaff, le biochimiste américain James Watson et le biologiste britannique Francis Crick proposèrent un modèle de double hélice aujourd'hui accepté comme correct (Watson and Crick, 1953).

Depuis ce temps, des progrès énormes ont été réalisés avec le séquençage de milliers de génome dont celui de l'être humain en 2001 pour les séquences brutes (Lander et al.,

2001) et 2004 pour le génome complet par le consortium international comptant 16 centres de séquençage à travers le monde (Human Genome Sequencing Consortium, 2004). Mais cela ne s'arrête pas là, il est également possible de nos jours d'étudier la structure tridimensionnelle de molécules (Wenk et al., 1992), de séquencer le transcriptome (Forrest and Carninci, 2009) (l'ARN étant connu depuis 1939 (Caspersson and Schultz, 1939)), d'étudier le protéome (Langen et al., 1997) mais également le métabolome (Oliver et al., 1998). Un monde s'ouvre à nous et de nombreux aspects restent à découvrir.

## I.2. L'expression génétique

Il est à présent entendu que le support de l'information génétique est l'ADN mais par quels mécanismes parvient-il à transmettre cette information pour qu'un organisme puisse fonctionner ? Ces mécanismes sont principalement la transcription et la traduction (Fig. 1) et constituent le dogme central de la biologie moléculaire formulé en 1958 par Crick (Crick, 1970).

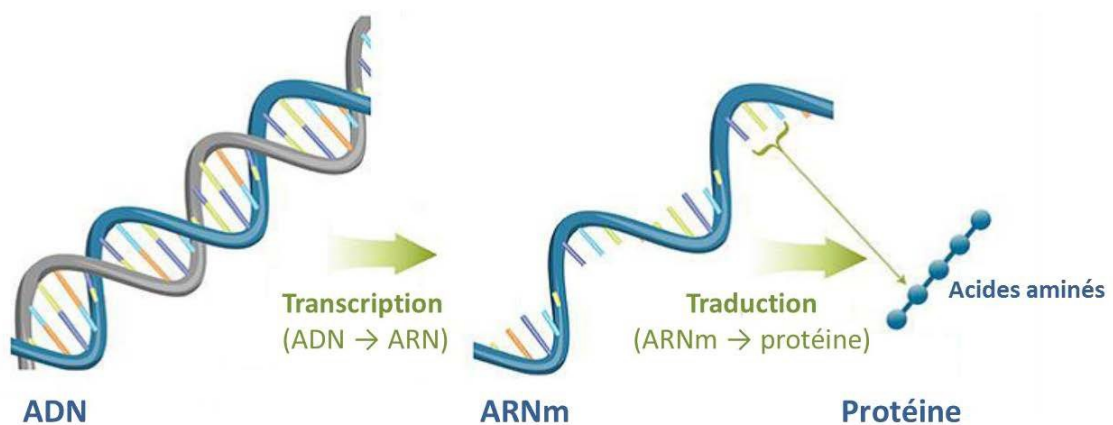


Figure 1. Principe du dogme central de la biologie moléculaire.  
(Adapté de [www.studynotes.ie](http://www.studynotes.ie))

L'information génétique encodée par l'ADN est transcrite par l'ARN polymérase en ARN messager (ARNm). L'ADN ne comportant que 4 nucléotides (nt) différents capables de s'apparier par des liaisons hydrogène pour conserver cette information, les gènes sont donc des combinaisons plus ou moins longues de ces



nucléotides : des bases puriques (Adénine et guanine, Tab.1) et des bases pyrimidiques (Thymine et cytosine, Tab. 2). L'adénine (A) se liant à la thymine (T) et la guanine (G) s'appariant à la cytosine (C). Par contre l'ARNm ne contient pas de thymine mais de l'uracile (U) qui ne diffère de la thymine que par l'absence d'un groupement méthyle sur le carbone 5 (Tab.2). La présence de ce groupement offre notamment une meilleure protection à l'ADN contre les dommages dus aux radiations et permet ainsi à ce dernier d'être plus stable face aux mutations (Lesk, 1969).

Tableau 1. Bases puriques

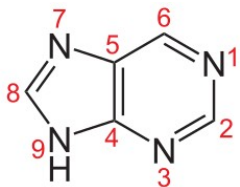
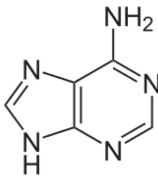
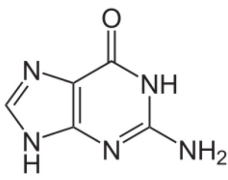
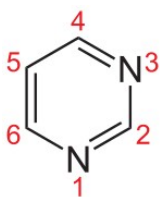
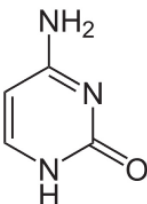
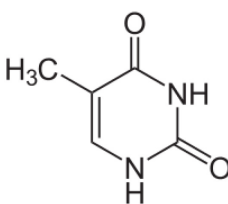
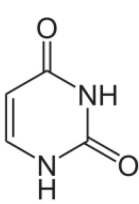
		
Purine	Adénine	Guanosine

Tableau 2. Bases pyrimidiques

			
Pyrimidine	Cytosine	Thymine	Uracile

L'ARNm une fois synthétisé va être reconnu par les ribosomes au niveau de leur site de liaison également appelé la séquence de Shine-Dalgarno (Shine and Dalgarno, 1974) qui vont alors lire sa séquence et regrouper les nucléotides trois par trois pour leur associer un acide aminé. Ce processus s'appelle la traduction et résulte en la synthèse de protéines. Il existe 64 manières différentes d'associer 3 nt (on parle alors de codon) mais seulement 20 acides aminés. Ce problème arithmétique est résolu par la redondance du code génétique qui permet à plusieurs codons différents d'être traduit en le même acide aminé. Le ribosome est constitué de deux sous-unités dont le poids moléculaire diffère entre procaryotes et eucaryotes. La petite sous-unité (30S chez les procaryotes et 40S chez les

eucaryotes) permet la lecture des codons de l'ARNm tandis que la grande sous-unité (50S chez les procaryotes et 60S chez les eucaryotes) assure la formation des liaisons peptidiques entre les acides aminés apportés par des ARN de transfert.

Les protéines sont donc des macromolécules produites lors de la traduction pour être les unités fonctionnelles des gènes et possèdent des fonctions très diverses. Nombre d'entre elles sont impliquées dans les fonctions structurales (membrane cellulaire, cytosquelette, etc.), motrices (actine, myosine, etc.), ou encore de transporteurs (ports membranaires, canaux, etc.). D'autres peuvent avoir des actions enzymatiques et vont pouvoir catalyser des réactions chimiques au sein du métabolisme cellulaire. D'autres encore vont être en charge de la signalisation cellulaire ou de la défense (hormones, anticorps, etc.) ou vont intervenir au niveau de l'expression des gènes (polymérases, facteurs de transcription, etc.).

Néanmoins, l'ARNm n'est pas le seul produit de la transcription. En effet, certains gènes vont encoder d'autres types d'ARN non-codants (ARNnc) car non reconnus par les ribosomes. Leurs fonctions sont très variées et parmi ces ARNnc, on peut citer des ARN de transfert (ARNt), les ARN ribosomiaux (ARNr) mais également de petits ARN qu'ils soient nucléaires (pARNn), nucléolaires (pARNno) ou les microARN (miARN).

L'existence de l'ARN semble logique dans le cas des eucaryotes où l'ADN est conservé au sein du noyau et ne peut passer par les pores nucléaires pour rejoindre le cytoplasme où se trouvent les ribosomes. Il y a donc nécessité d'une molécule servant d'intermédiaire entre le noyau et le cytoplasme. Pourtant chez les procaryotes, il y a présence d'ARN même s'il n'y a pas de noyau. Toutefois, il existe une théorie selon laquelle l'ARN serait une molécule bien plus vieille que l'ADN et aurait servi pour le stockage de l'information génétique. De plus des molécules d'ARN peuvent avoir des fonctions enzymatiques et auraient pu donc se suffire à elles-mêmes dans un monde primordial (Robertson and Joyce, 2012).

## II. Régulation génétique bactérienne

L'ADN encode pour un vaste panel de gènes dont l'organisme dispose afin de répondre au mieux aux conditions environnementales et assurer ainsi sa survie. Bien évidemment, l'organisme ou la cellule n'expriment pas en permanence la totalité des gènes dont ils disposent. Ce serait une bien trop grande charge énergétique et puisque c'est une ressource limitée, ce serait contre-productif. En effet, certains gènes sont dévolus à des situations précises comme le choc thermique, le transfert de gènes ou de d'éléments génétiques mobiles comme des plasmides ou encore la mort cellulaire.

Pour que la cellule soit en mesure d'exprimer le bon gène au bon moment et assurer une croissance optimale et un bon fonctionnement quelles que soient les conditions environnementales, elle a besoin de réguler son expression génétique et donc à fortiori son métabolisme. Or parfois cette régulation doit être très fine pour que la cellule demeure compétitive ou simplement survive. Pour cela, elle dispose de plusieurs mécanismes dont les facteurs de transcriptions, les facteurs sigma, les facteurs d'élongation de la transcription, les facteurs de terminaison (ou d'anti-terminaison), les ARN régulateurs ainsi que de protéines telles les ribonucléases. Ces acteurs vont pouvoir intervenir à différents niveaux et vont être plus amplement développés ci-après de manière non exhaustive (Fig. 2).

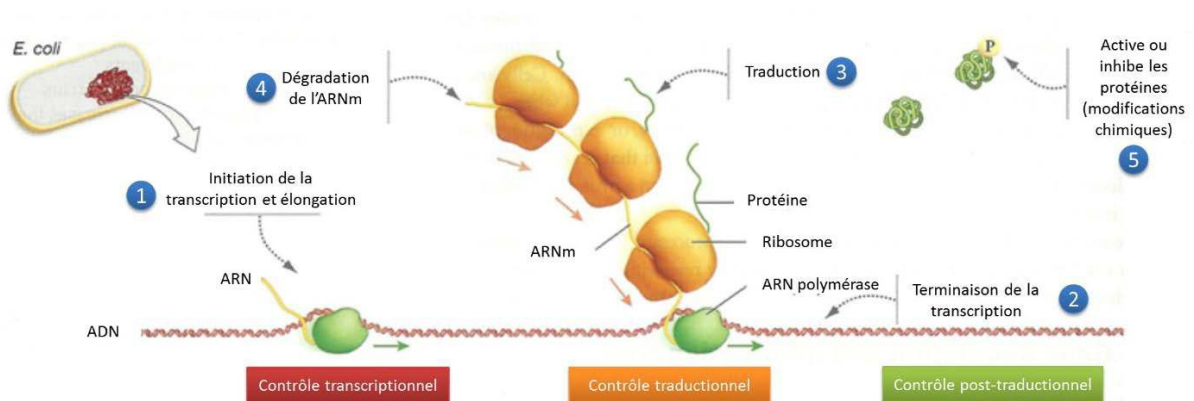


Figure 2. Schématisation des possibilités principales de régulation de l'expression génétique chez la bactérie. (Adapté de [discoveryandinnovation.com](http://discoveryandinnovation.com))

## II.1. Régulation de la transcription

La transcription étant la première étape de l'expression des gènes, il est important pour la bactérie de pouvoir la réguler de manière efficace afin d'éviter de produire inutilement des ARN et donc de gaspiller de l'énergie. L'initiation de la transcription (1 sur la Fig. 2) est donc une étape clef et va dépendre de facteurs de transcription. Ils sont très nombreux chez les bactéries (plus de 200 par exemple chez *B. subtilis*) et comptent parmi leurs rangs les facteurs sigma. Mais chaque étape peut être régulée et a son importance.

### II.1.1. Les facteurs sigma

La bactérie ne dispose pas comme les organismes eucaryotes de plusieurs ARN polymérases. Pour pouvoir distinguer les gènes à transcrire, elle va utiliser différents facteurs sigma ( $\sigma$ ). En effet, l'ARN polymérase est constituée de plusieurs sous-unités pour former une holoenzyme  $\alpha_2\beta\beta'\omega$  à laquelle est ajouté un  $\sigma$ . Les fonctions de ces différentes sous-unités sont bien définies : 2  $\alpha$  permettant l'assemblage du cœur ( $\alpha$ -NTD) et la liaison non-spécifique au promoteur ( $\alpha$ -CTD) ; 1 sous-unité  $\beta$  dotée de l'activité catalytique nucléotidyl-transférase ; 1 sous-unité  $\beta'$  liant l'ADN ; 1 sous-unité  $\omega$  permettant l'assemblage et la stabilité du cœur et enfin la sous-unité  $\sigma$  permettant la liaison au promoteur et la formation de la boucle de transcription (Campbell et al., 2002; Vassylyev et al., 2002). Il en existe deux familles phylogénétiquement distinctes au fonctionnement différent et identifiées par leur poids moléculaire : les 70 qui sont ubiquitaires et ne requièrent pas d'activateur pour ouvrir l'ADN et les 54 qui eux requièrent un activateur et sont présents dans environ 65% des bactéries. Parmi les  $\sigma^{70}$ , on dénombre une vingtaine de membres dont quelques exemples sont donnés dans le tableau 3 sans oublier les  $\sigma^{E, F, H, K}$  impliqués à différentes étapes de la sporulation de certains organismes (*B. subtilis* notamment) (Kroos et al., 1999).

Ces facteurs vont rentrer en compétition entre eux pour les ARN polymérases et plusieurs paramètres entrent en jeu pour déterminer le vainqueur. Tout d'abord, leur concentration relative (synthèse *de novo*, modifications post-traductionnelles ou dégradation protéolytique) mais également la présence de facteurs anti-sigma pouvant les

séquestrer, de molécules de régulation et finalement la capacité des facteurs  $\sigma$  à reconnaître des promoteurs dégénérés.

Tableau 3. Exemples de facteurs sigma de la famille 70.

Nom	Séquence de reconnue	Fonction
$\sigma^{70}$ RpoD	TTGACA	Plupart des gènes, facteur principal pour croissance normale, « housekeeping ».
$\sigma^{38}$ RpoS	CCGGCG	Phase stationnaire, stress oxydatif et osmotique.
$\sigma^{32}$ RpoH	TNTCNCCTTGAA	Réponse aux chocs thermiques.
$\sigma^{28}$ FliA	TAAA	Gènes impliqués dans la synthèse des flagelles.
$\sigma^{24}$ RpoE	GAACCTT	Réponse aux protéines mal repliées dans le périplasme.
$\sigma^{19}$ Fecl	AAGGAAAAT	Certains gènes impliqués dans le transport du fer.

### II.1.2. Réactions enzymatiques

Si une bactérie peut réguler l'initiation de la transcription via des séquences promotrices particulières, elle peut également interagir sur la concentration intracellulaire de certaines enzymes inductibles, que ce soit à la hausse ou à la baisse. Il s'agit d'un mécanisme opérant en réponse à la présence ou non d'un certain ligand. Un exemple bien connu de ce processus est donné par l'opéron lactose (*lac*) impliqué dans le transport et la synthèse du lactose chez plusieurs bactéries du microbiote intestinal et ce en réponse à la disponibilité en glucose et en lactose. Cet opéron comporte 3 gènes structurels : *lacZ*, *lacY* et *lacA* présents dans le même sens, immédiatement adjacents et transcrits sur un même ARNm polycistronique. Cet opéron confère la possibilité de digérer efficacement le lactose en glucose et en galactose via l'enzyme  $\beta$ -galactosidase encodée par le gène *lacZ*. Cette réaction, source d'énergie pour la cellule, n'est toutefois pas nécessaire lorsqu'il n'y a pas de lactose disponible dans le milieu ou s'il y existe une source d'énergie plus rentable. La bactérie doit donc réguler la production de cette enzyme et possède pour cela un mécanisme en deux composantes.

En l'absence de lactose, l'expression de l'opéron *lac* est enrayée par le répresseur *lac* (*lacI*) qui vient se lier très fortement à la séquence opératrice située en amont du gène *lacZ* empêchant ainsi la liaison de l'ARN polymérase sur le promoteur et donc la transcription des gènes. Cette inhibition peut être levée par l'allolactose, un isomère du lactose qui entraîne par sa liaison au répresseur un changement de conformation de ce dernier et inhibe son interaction avec la séquence opératrice.

Il est également possible à la cellule de réguler l'expression de l'opéron en réponse à la concentration en glucose. En effet, la protéine activatrice des catabolites (CAP) est alors activée et se lie à l'adénosine monophosphate cyclique (AMPC). Le complexe ainsi formé se lie à son tour sur une séquence en amont du promoteur et aide l'ARN polymérase à venir se fixer sur l'ADN pour permettre la production de  $\beta$ -galactosidase.

Il existe également des opérons dont l'expression est répressible comme celui de l'opéron tryptophane codant pour 5 gènes impliqués dans la synthèse de cet acide aminé dont la tryptophane synthase. Le tryptophane agit comme un répresseur et vient se lier sur une séquence inhibitrice, donnant ainsi un rétro-contrôle négatif sur sa biosynthèse.

### II.1.3. Élongation

Au cours de la transcription, durant l'étape d'élongation, les ARN sont également régulés. En effet, ils se replient en des structures secondaires élaborées pour s'associer avec des protéines ribosomales et être traduits. Mais des signaux d'acides nucléiques et des protéines régulatrices vont venir moduler ces événements. Pour cela, ils vont notamment indiquer à l'ARN polymérase où effectuer une pause, où se termine la transcription ou encore agir comme des obstacles au cheminement de l'enzyme (Belogurov and Artsimovitch, 2015). L'élongation est faite par un complexe d'élongation (EC) constitué de l'ARN polymérase, de l'ADN partiellement ouvert et de l'ARN transcrit naissant.

La transcription peut être interrompue provisoirement par ce qu'on appelle des pauses. Ce sont des mécanismes par lesquels l'ARN polymérase va s'isomériser et ralentir son incorporation de nucléotides entre 10 et 1000 fois, et ce même *in vitro*. Pauser est une propriété intrinsèque de l'ARN polymérase dont la probabilité est augmentée par certaines séquences au sein de la matrice d'ADN. Cette isomérisation de l'ARN polymérase peut entraîner des réarrangements structuraux qui stabilisent et donc allongent la pause.

Néanmoins, ce ne sont pas les seuls événements qui peuvent interférer avec l'élongation d'un ARN. Il est également possible à l'ARN polymérase de revenir sur ses pas bien que ce phénomène soit en général empêché par la présence de ribosomes, de s'arrêter en cours de transcription ou bien évidemment de terminer la transcription sans

encombre. Ceci est rendu possible par des facteurs d'élongation ou encore par des systèmes d'anti-termination comme les protéines ribonucléiques Nus. NusA, la plus importante en termes de taille et la pièce centrale de ce système, comporte plusieurs domaines spécialisés et intervient dans l'augmentation de la sensibilité de l'ARN polymérase aux tiges-boucles qu'il peut former et qui sont impliquées dans les pauses et les terminaisons intrinsèques. D'autre part, elle intervient dans le repliement co-transcriptionnel de l'ARN et enfin participe à l'assemblage des complexes d'anti-termination.

#### II.1.4. Termination

La terminaison est comme son nom l'indique la dernière étape de la transcription et la garante d'une expression génétique correcte tout en permettant la libération de l'ARN polymérase. Deux grands mécanismes permettent d'arriver à ce résultat chez les bactéries : la terminaison intrinsèque et la terminaison rho-dépendante (Ray-Soni et al., 2016).

La terminaison intrinsèque consiste en la reconnaissance de séquences particulières sur l'ARN : une séquence dite canonique riche en G-C à la fin du gène, constituant la tige terminatrice, suivie immédiatement par une queue poly-U de 7 à 8 nucléotides. Lorsque l'ARN polymérase rencontre la fin de cette queue poly-U, elle va faire une pause à son extrémité 3' ce qui va permettre à la tige terminatrice de se former. Cette phase de pause est très importante pour la terminaison et rend accessibles les nucléotides impliqués dans la tige (la polymérase en protège 14 en son sein). La formation de la tige induit la dissociation de l'hybride ADN/ARN 3 à 4 nt en amont et déstabilise le complexe d'élongation.

La terminaison est également possible par l'intermédiaire de translocases d'ARN : les protéines Rho qui sont des hélicases homohexamériques de la famille RecA, famille jouant un rôle crucial dans la recombinaison génétique et la réparation des cassures double-brins de l'ADN (Leipe et al., 2000). Les protéines Rho reconnaissent de manière non spécifique des signaux sur le transcrit d'ARN naissant, qui sont généralement des séquences riches en C et en pauvres en G pour éviter la formation de structures secondaires. La reconnaissance au brin d'ARN se fait au niveau de sites *rut* (*Rho-utilization site*) se situant en amont du site de terminaison et faisant entre 80 et 90 nt de

long. La longue séquence de ces sites permet ainsi à l'ARN de se lier à chacun des monomères de la protéine. Cette première reconnaissance est suivie par l'enroulement du transcrit au sein de la protéine Rho via son pore central. Ce processus de « rembobinage » est appelé translocation et s'effectue par hydrolyse de l'ATP. Une fois celle-ci achevée par la translocation de la totalité de l'ARN en amont du complexe d'élongation à travers le pore de la protéine Rho, l'EC est alors déstabilisé et l'ARN libéré.

## II.2. Régulation de la traduction

L'intérêt de réguler la traduction au lieu de la transcription peut sembler une perte d'énergie pour la cellule car l'ARNm est produit de toute manière (Étape 3 sur la Fig. 2). Néanmoins, cela permet à celle-ci d'avoir des temps de réaction plus courts face à des stimuli environnementaux. Et ce, même si les délais sont moins importants chez une bactérie que pour une cellule eucaryote qui doit potentiellement épisser, coiffer, polyadényler et exporter l'ARNm dans le cytoplasme. Ici, au lieu de passer par l'étape de transcription, il ne reste qu'à traduire l'ARNm. Cependant, si aucun stimulus ne parvient à la cellule avant que l'ARNm ne soit dégradé, cette synthèse aura effectivement été vaine. Si la Nature a sélectionné ce mécanisme, c'est bien parce que le rapport coût/gain était avantageux pour l'organisme. Comme pour la transcription, la régulation peut s'effectuer à plusieurs niveaux.

### II.2.1. Initiation de la traduction

L'étape de la traduction la plus régulée est l'initiation même si plus rarement elle peut être modulée pendant l'élongation ou la terminaison. Une manière simple de réguler la traduction est de contrôler le nombre de ribosomes présents au sein de la cellule, surtout que bon nombre de protéines ribosomiques sont elles-mêmes régulées au niveau traductionnel chez *Escherichia coli* par exemple. D'autre part pour initier la traduction, la sous-unité 30S du ribosome doit se fixer sur le codon de démarrage via une séquence de reconnaissance en amont, la séquence de Shine-Dalgarno. De ce fait, si cette séquence est séquestrée par une structure secondaire de l'ARNm ou par la présence d'une autre protéine régulatrice, le ribosome ne pourra pas initier la traduction. Il s'agit ici d'une



régulation en *cis*, c'est-à-dire agissant dans son entourage immédiat. Il est également possible d'observer une régulation en *trans* et de masquer le site d'initiation par un ARN anti-sens qui va former un ARN double brin qui là aussi empêchera la liaison du ribosome. C'est le cas par exemple pour la porine OmpF dont l'ARNm peut être lié par un ARN antisens (MicF) pour une inhibition de son expression (Ziervogel and Roux, 2013). Une autre stratégie de régulation consiste à décaler du cadre de lecture, le codon de démarrage se trouvant dans un cadre et la terminaison dans un autre, décalé d'un nucléotide en amont ou en aval. Le ribosome est alors obligé de glisser d'un cadre à l'autre via une structure secondaire stable d'ARN qui va pauser le ribosome et permettre le glissement. C'est un mécanisme notamment utilisé par les bactéries pour moduler la concentration de facteurs de terminaison de la traduction.

Les procaryotes disposent également de facteurs d'initiation (PIF-1 à PIF-3) qui s'associent à la sous-unité 30S du ribosome. Au niveau du site A pour empêcher l'ARNt chargé d'entrer pour PIF-1 et au niveau du site P pour permettre la formation de la sous-unité 70S pour PIF-2. PIF-3 quant à lui est moins retrouvé chez les bactéries mais est nécessaire à la liaison de la sous-unité 30S au site d'initiation de la traduction sur l'ARNm, il devra cependant être libéré pour que la sous-unité 50S puisse venir se fixer au complexe d'initiation.

#### II.2.2. Modification de la stabilité des ARNm

Une autre manière simple de réguler la traduction et de dégrader l'ARNm par l'action de ribonucléases avant que celui-ci ne soit transcrit ou au contraire de le protéger via la liaison de protéines jusqu'à ce que les ribosomes puissent le prendre en charge (Étape 4 sur la Fig. 2) (Roth and Breaker, 2009).

Ce résultat peut également être obtenu via l'interaction d'un ARN anti-sens et d'un ARNm dont la structure secondaire va alors être modifiée. Cette modification structurale peut avoir de grandes répercussions sur la demie vie de l'ARNm et sa stabilité. Un exemple en est donné par l'ARN anti-sens OOP du phage  $\lambda$  qui est complémentaire de l'extrémité 3' d'un autre ARNm du phage (le répresseur cII). Leur association va mener au clivage des deux ARN (Georg and Hess, 2011).

### II.2.3. Modification ou dissociation des ribosomes

Il existe des facteurs dont le rôle est de dissocier le ribosome, comme la GTPase HflX chez *E. coli*. Ce facteur intervient en cas de choc thermique et peut dissocier aussi bien les ribosomes libres que ceux associés à des ARNm. Pour cela, le domaine N-terminal de HflX se fixe au niveau de la peptidyl-transférase du ribosome et induit des changements conformationnels importants entre les deux sous-unités du ribosome les menant à leur dissociation (Zhang et al., 2015). Au lieu de dissocier un ribosome déjà formé, il est également possible d'empêcher cette formation. C'est l'action de la protéine RsfA retrouvée dans la grande majorité des eubactéries et qui se lie à une protéine de la grande sous-unité (L14) pour empêcher les deux sous-unités de s'assembler et de former un ribosome fonctionnel. Ceci résulte en un ralentissement voire à un arrêt de la traduction par le ribosome ciblé (Häuser et al., 2012).

Lorsque la bactérie entre en phase stationnaire, elle a tout intérêt à limiter sa production de protéines en raison du peu de nutriments disponibles. Cette limitation intervient par différents mécanismes comme la polymérisation des ribosomes 70S en dimères 90S lorsqu'ils lient un facteur de modulation ribosomal (RMF). Ces dimères peuvent lier à leur tour un facteur d'hibernation (HPF) et forment un complexe de 100S qui est inactif, en hibernation (Kato et al., 2010).

### II.3. Régulation métabolique (post-traductionnelle)

Bien que moins employée, il est également possible d'agir au niveau métabolique (Fig. 2, étape 5) en modifiant les enzymes synthétisées ou en venant agir via une boucle de rétro-inhibition sur une enzyme intervenant en amont d'une chaîne métabolique. Ainsi le métabolite final peut venir inhiber une étape de sa biosynthèse si jamais il devait être présent en quantités trop importantes. L'inhibition, ne touchant que l'enzyme visée et épargnant les autres étapes, serait alors levée quand le métabolite est en concentration trop faible.

## II.4. Antibiotiques et résistances

La cellule a donc tout un arsenal pour réguler l'expression de ses gènes. Toutefois, nous pouvons aussi jouer à ce jeu. Bien que des traitements contre les infections existent depuis l'Antiquité, on considère l'ère des antibiotiques comme ayant débuté en 1928 avec la découverte de la pénicilline par Sir Alexander Fleming. Ce fut un tournant majeur dans le traitement des maladies et a depuis permis de sauver des millions de vies. Mais c'est surtout à partir des années 40 qu'en Occident s'est développée cette ère, et ce notamment pour le traitement des infections pendant la seconde guerre mondiale. Ces traitements presque miraculeux étaient malheureusement trop beaux pour durer. En effet, dès les années 50, les premières résistances furent constatées chez les bactéries. Une tendance qui s'est amplifiée depuis et largement accélérée (Ventola, 2015). Il existe une théorie en écologie qui illustre assez bien ce qui se passe de nos jours. Il s'agit de l'hypothèse de la Reine rouge (*Red Queen hypothesis*) par Leigh Van Valen qui se base sur la Reine Rouge du livre « De l'autre côté du miroir », suite d'Alice au Pays des Merveilles de Lewis Carroll (Van Valen, 1973). Cette image repose sur la course constante pour rester à la même place dans une course à l'armement entre nous d'un côté et les microorganismes de l'autre. Nous développons des antibiotiques pour prendre l'avantage et peu de temps après apparaît une résistance et tout est à refaire. Actuellement le problème est que nous perdons cette course. De plus en plus de microorganismes deviennent résistants alors qu'il y a de moins en moins de nouveaux antibiotiques sur le marché (Fig. 3).

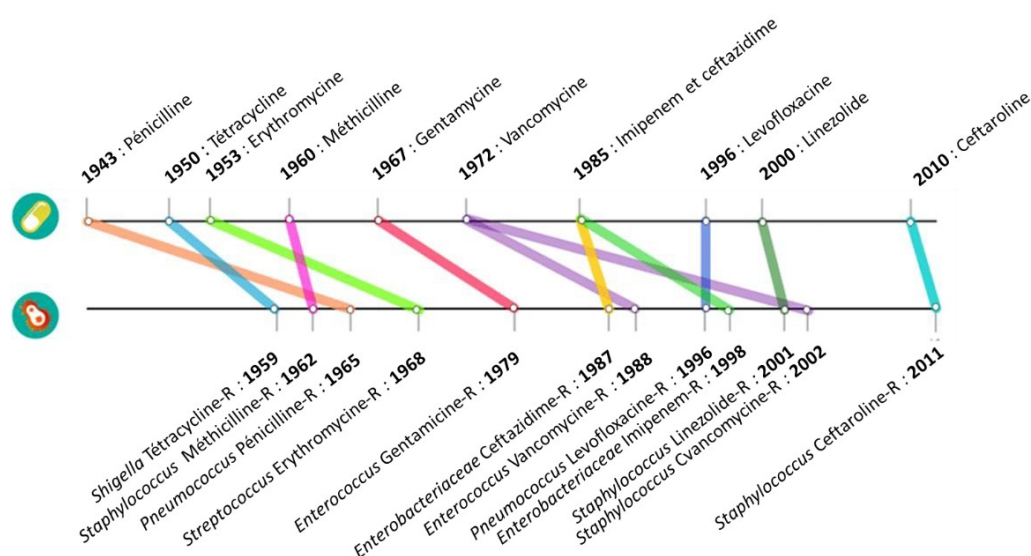


Figure 3. Chronologie de la mise sur le marché de quelques antibiotiques (haut) et l'apparition de la première résistance à celui-ci (bas). Données CDC.

La résistance des microorganismes aux antibiotiques n'est pourtant pas un phénomène nouveau. Les gènes codant pour ces résistances sont retrouvés dans la nature mais généralement dans d'autres microorganismes que ceux ciblés initialement par l'antibiotique. Le problème est que les bactéries sont capables de se transférer des gènes les unes aux autres par des transferts de gènes horizontaux par exemple. Ainsi une bactérie pathogène peut très bien acquérir une résistance conférée par une bactérie tout à fait inoffensive pour l'Homme. Quelques-unes des cibles les plus communes des antibiotiques ainsi que les stratégies de résistance bactériennes sont présentées en figure 4.

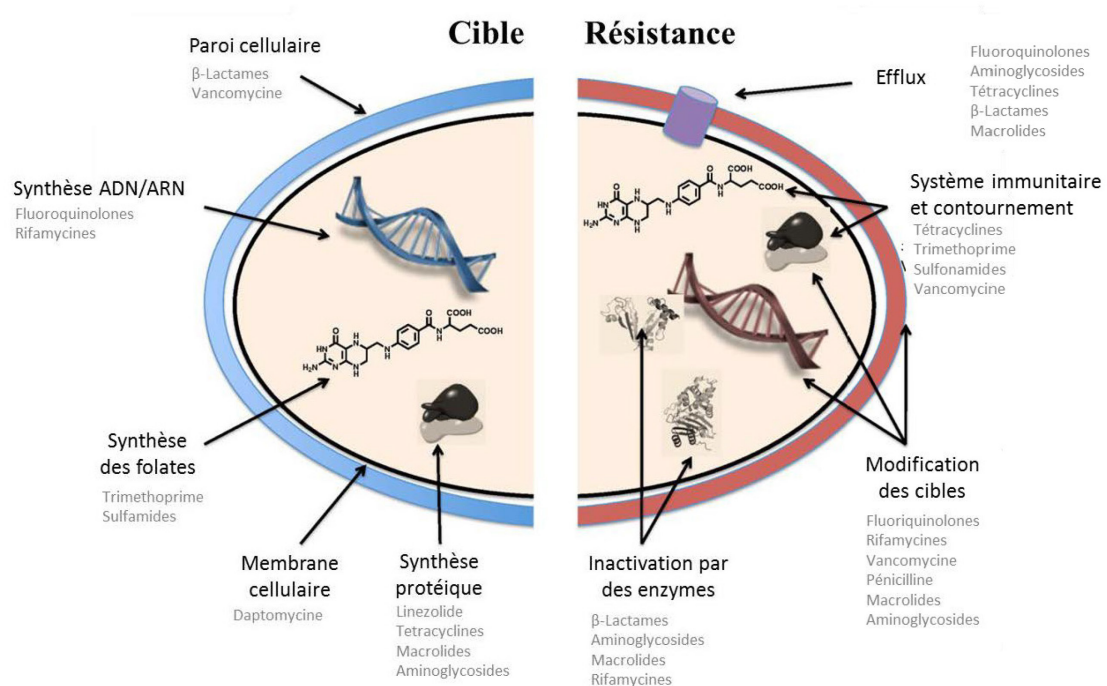


Figure 4. Cibles les plus communes des antibiotiques et stratégies de résistances. Adapté de (Wright, 2010).

Nous avons également notre part de responsabilité dans l'apparition de plus en plus rapide de ces résistances. On peut citer par exemple les abus comme les mauvaises utilisations, l'augmentation de leur utilisation dans les élevages de bétail ainsi que le peu de nouveaux antibiotiques découverts (Michael et al., 2014). Il est maintenant notoire depuis quelques années que des antibiotiques ont été prescrits à tort et à travers dès qu'une infection était suspectée. C'était à priori plus simple pour tout le monde de prescrire des antibiotiques à large spectre que d'essayer d'identifier la bactérie responsable de l'infection. Cet abus d'antibiotiques a considérablement permis l'apparition de

résistances. En effet, en donnant des antibiothérapies, on va sélectionner les espèces qui y sont résistantes et leur permettre de proliférer là où elles n'auraient peut-être pas été en mesure de le faire en raison de la compétition entre les espèces dans un microbiote. Depuis quelques années, des efforts sont faits afin de réduire les prescriptions aux cas où ils sont nécessaires et faire comprendre au public que certaines infections, comme celles d'origines virales, ne nécessitent pas d'antibiotiques. On peut citer notamment la campagne française lancée en 2002 sur la sensibilisation du grand public avec des spots publicitaires aux slogans tels que « Les antibiotiques, c'est pas automatique » ou « Les antibiotiques, utilisés à tort, ils deviendront moins forts ». L'efficacité de cette campagne, évaluée sur 5 ans, a permis entre 2002 et 2007 de diminuer les prescriptions d'antibiotiques de 26,5% (Sabuncu et al., 2009).

Néanmoins, le problème des résistances demeure et est dans le cas des antibiotiques, inévitable. Les efforts entrepris pour limiter l'utilisation des antibiotiques ne servent qu'à retarder l'échéance mais ces efforts sont essentiels. Dès 1945, Fleming s'inquiétait de cet engouement et prédisait « le public va demander le médicament ... et alors commencera une ère ... d'abus ». En 2015, l'OMS a déclaré que nous allions entrer dans une ère post-antibiotiques avec le retour de maladies infectieuses que nous ne pourrions plus traiter faute d'antibiotiques efficaces. Ainsi, plusieurs cas de bactéries résistantes à tous les antibiotiques connus ont été recensés dans les hôpitaux. Le dernier cas reporté par le CDC a coûté la vie en août 2016 à une américaine de 70 ans infectée par un *Klebsiella pneumoniae* résistant aux 26 antibiotiques disponibles (Chen et al., 2017). C'est pourquoi il est de nos jours nécessaire de trouver des thérapies alternatives aux antibiotiques qui idéalement n'induisent pas ou peu d'apparition de résistances. C'est notamment le cas des riborégulateurs.

### III. Riborégulateurs

#### III.1. Historique

Les riborégulateurs (en anglais « *riboswitches* ») sont des structures d'ARN retrouvées généralement dans la portion 5'-non codante (5'-UTR) des ARNm (Fig. 5). Identifiés pour la première en 1998 (Grundy and Henkin, 1998), ce n'est qu'en 2002 qu'ils seront décrits (Nahvi et al., 2002). Le terme de « *riboswitch* » sera quant à lui introduit en 2002 par l'équipe de Ronald R. Breaker pour qualifier cette régulation *en cis* par une portion d'ARNm (Winkler et al., 2002). Ils brevètent la découverte un an plus tard (US 7794931 B2).

Cependant au début des années 90, il était déjà su que des ARN étaient en mesure de lier de petites protéines ainsi que de petits ligands (Ellington and Szostak, 1990). Moins de 10 ans plus tard, il était démontré par exemple que la région 5'-UTR pouvait avoir un impact sur l'expression du gène en aval. Des mutations introduites dans cette région chez *B. subtilis* et *Bacillus amyloliquefaciens* en amont de gènes associés à la synthèse de riboflavine avaient un impact sur celle-ci. Or cette région d'ARNm formait une structure complexe. C'est ainsi que l'inhibition des gènes des vitamines B1, B2 et B12 par la thiamine, la riboflavine et la cobalamine fut élucidée (Gelfand et al., 1999). Il fut alors démontré que la liaison du ligand permettait de stabiliser la conformation d'une partie hautement conservée de ces ARN tout en induisant un changement conformationnel dans une partie, elle, très variable. La combinaison de ces deux structures d'ARN affectant l'expression du gène en aval, d'où le nom de riborégulateur.

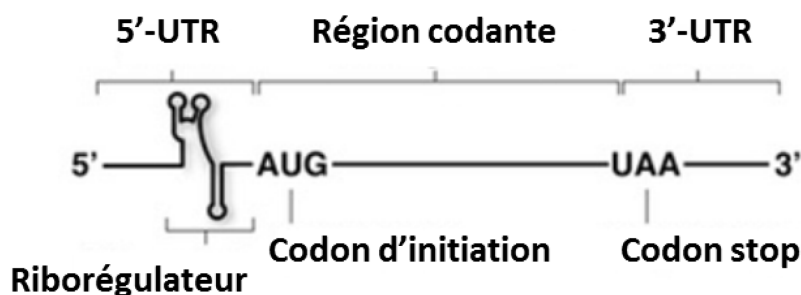


Figure 5. Schéma de la structure d'un ARNm bactérien doté d'un riborégulateur.

## III.2. Anatomie et fonctionnement

Typiquement, un riborégulateur est constitué de deux régions : un aptamère capable de lier un ligand spécifique, et une plateforme d'expression pouvant influencer, par des changements de conformation de l'ARNm, sur l'expression du gène situé en aval (Fig. 5).

Ils sont retrouvés préférentiellement chez les bactéries, 4% du génome de *B. subtilis* est ainsi sous la dépendance de riborégulateurs (Mandal et al., 2003). Toutefois on en retrouve également chez les plantes (Bocobza et al., 2007; Wachter et al., 2007), les champignons filamenteux (Cheah et al., 2007; Galagan et al., 2005), les phages (Ramirez, 2016; Sudarsan et al., 2008) et les archées (Weinberg et al., 2010). Dans le cas des plantes et des champignons, il s'agit des deux seuls riborégulateurs (TPP et *crcB*) présents jusqu'à maintenant chez des eucaryotes. Leur mécanisme diffère car ils sont présents en 5' ou en 3'-UTR selon les espèces et régulent l'épissage alternatif du gène (Wachter et al., 2007). Dans l'état actuel des connaissances, aucun riborégulateur n'a été identifié chez les mammifères.

Les mécanismes les plus couramment retrouvés concernent la régulation de la transcription ou de la traduction en *cis* de gènes impliqués dans la biosynthèse ou le transport de métabolites. Néanmoins, il en existe d'autres comme la dégradation de l'ARNm ou son épissage alternatif. Généralement, l'absence de ligand se solde par une inactivation de l'expression du gène, on parle alors d'une conformation « OFF » tandis que la présence du ligand au contraire permet l'expression du gène (conformation « ON »).

### III.2.1. Régulation transcriptionnelle

La liaison du ligand permet de stabiliser la conformation du riborégulateur et empêche (Fig. 6, partie gauche en orange) la formation d'une boucle anti-terminatrice qui serait complémentaire de la partie 3' de la tige P1 du riborégulateur (figurée en bleu). Ceci mène à la formation d'une boucle de terminaison rho-indépendante alternative (haut) ou rho-dépendante (bas) qui résultent toutes deux en une terminaison prématurée de la transcription par déstabilisation de l'ARN polymérase du complexe ADN/ARNm. L'ARNm n'étant pas transcrit, le gène n'est alors pas exprimé.



Il arrive néanmoins plus rarement que la liaison du métabolite entraîne la stabilisation d'une boucle anti-terminatrice qui elle va permettre à l'ARN polymérase de poursuivre la transcription et de donner un ARNm fonctionnel qui pourra être traduit (Fig.6 centre en vert).

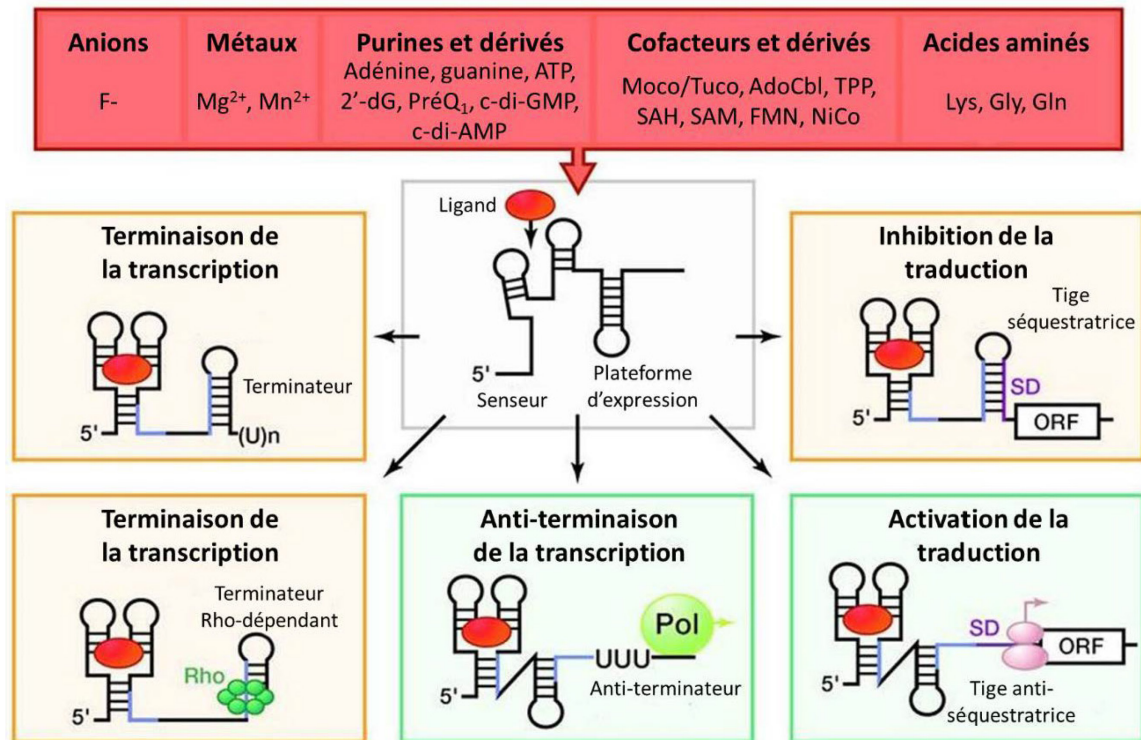


Figure 6. Différents mécanismes utilisés pour la régulation de la transcription et de la traduction par les riborégulateurs (adapté de (Serganov and Nudler, 2013)).

### III.2.2. Régulation traductionnelle

La régulation de la traduction va intervenir au niveau du site de liaison du ribosome (RBS) ou séquence Shine-Dalgarno (SD, en violet sur la Fig. 6) sur un ARNm complètement transcrit. L'expression des gènes visualisée par le cadre de lecture ouvert (ORF pour *open reading frame*) peut donc être inhibée dans le cas où la séquence SD est séquestrée ce qui empêche l'initiation de la traduction. Par contre, il est également possible d'obtenir l'effet inverse suite à la liaison du ligand, c'est-à-dire la formation d'une tige anti-séquestratrice qui rend accessible la séquence SD au ribosome (Fig.6 droite).



### III.2.3. Régulation par dégradation (ribozyme)

Un cas plus rare, retrouvé chez les bactéries Gram-positives, est le clivage de l'ARNm par une RNase (Fig. 7). Ainsi, la liaison de la glucosamine-6-phosphate (GlnN6P) entraîne le clivage par le ribosome/ribozyme *glmS* de son ARNm en 5' (Klein and Ferré-D'Amaré, 2006). L'extrémité 5'-OH libérée va alors être reconnue par la RNase J et l'ARNm est ensuite dégradé par la machinerie cellulaire. La liaison du GlnN6P n'entraînant pas de changement de conformation majeur, le ribozyme *glmS* serait préorganisé pour lier son métabolite ce qui suggère que celui-ci agirait comme un cofacteur du clivage des liaisons phosphodiester. L'avantage de la RNase J est que contrairement à celle d'*E. coli* (la RNase E impliquée dans le dégradosome du riborégulateur lysine (Caron et al., 2012)), elle est parfaitement en mesure de dégrader *in vivo* les transcrits présentant des groupements hydroxyles en 5'. Un autre exemple est le riborégulateur liant le c-di-GMP (Smith et al., 2009).

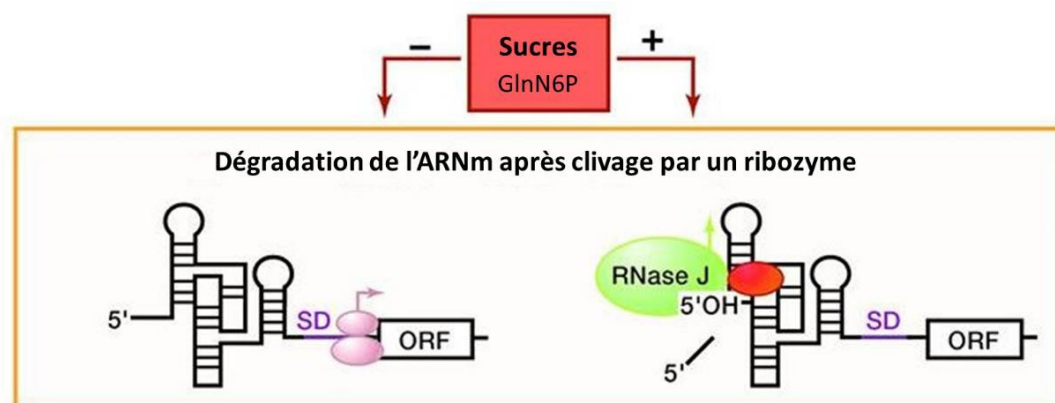


Figure 7. Mécanisme de régulation par clivage de l'ARNm. Adapté de (Serganov and Nudler, 2013).

### III.2.4. Régulation de l'épissage alternatif

Ce mode de régulation est intéressant car il est le seul retrouvé à ce jour à la fois chez les procaryotes et chez les eucaryotes. Dans l'exemple donné dans la figure 8, l'épissage alternatif a lieu chez *Clostridium difficile* où le c-di-GMP se lie à un riborégulateur ayant également des fonctions de ribozyme. Fonctions qui combinent l'épissage de l'ARNm produit ainsi que l'activation de la traduction (Hartig, 2010). À gauche, l'absence de ligand entraîne l'utilisation du second site de clivage du GTP (GTP<sub>2</sub>) et la formation d'un

ARNm avec une séquence SD tronquée qui n'est donc pas traduit. Au contraire, la présence de ligand (droite) associée à celle de GTP mène à l'utilisation du premier site de clivage du GTP (GTP<sub>1</sub>) et à l'autoexcision (en vert) d'un intron de groupe I. Cette excision est rendue possible via les sites de clivage spécifiques représentés par des ronds verts et il en résulte la réunion des deux parties de la séquence SD et à fortiori à la traduction de l'ARNm.

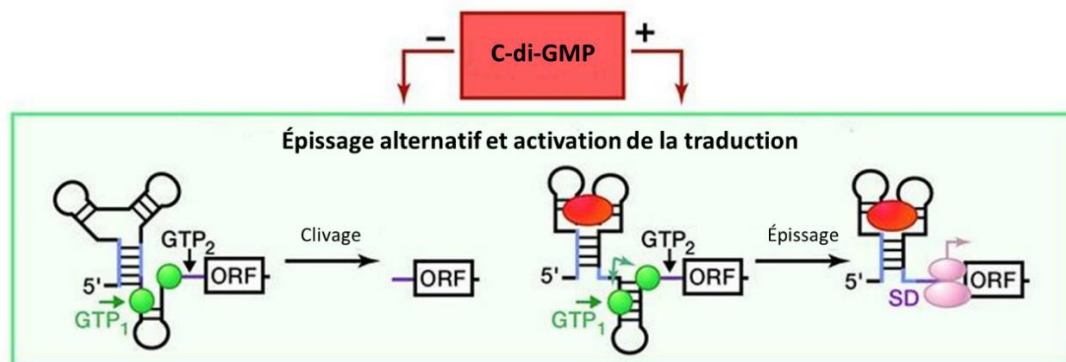


Figure 8. Mécanisme de régulation par épissage alternatif chez *C. difficile*. Adapté de (Serganov and Nudler, 2013).

Chez les eucaryotes, le riborégulateur le plus connu est celui liant le TPP (Cheah et al., 2007) et est retrouvé à l'extrémité 3' de l'ARNm qu'il régule. Il peut ainsi contrôler l'accès à des sites d'épissage alternatif. Le mode de régulation va dépendre de l'organisme considéré. Ainsi chez les champignons, l'absence de TPP se solde par la formation de structures secondaires qui promeuvent l'utilisation de sites d'épissage distaux et donc l'élimination des régions se trouvant entre ces sites. L'ARNm est alors produit dans son ensemble et le gène exprimé. En présence de ligand, on observe la stabilisation du riborégulateur ce qui entraîne l'utilisation d'autres sites de clivage qui n'étaient pas accessibles en absence de TPP. Le produit de cet épissage alternatif ne contient qu'une fraction de l'ORF ( $\mu$ ORF) et le gène ne peut être traduit. Chez les algues, la régulation porte sur l'élimination d'un codon stop dans la séquence du riborégulateur en absence de ligand qui autrement entraîne la terminaison prématurée de la traduction. Pour les plantes supérieures, le clivage influe sur les modifications post-traductionnelles que peut subir l'ARNm. Ainsi en absence de ligand, les sites de clivages entourant le signal de polyadénylation ne sont pas accessibles et l'ARNm est alors polyadénylé ce qui a pour conséquence l'obtention d'un ARNm stable avec une courte extrémité 3'. La

liaison du ligand va rendre accessibles ces sites, la séquence signal sera alors épissée ce qui résultera en un ARNm polyadénylé et pourvu d'une longue extrémité 3' le rendant plus instable.

#### III.2.5. Régulation anti-sens

Un autre mécanisme particulier est donné par la régulation anti-sens. En effet, les exemples donnés précédemment faisaient état d'une régulation en *cis* de l'expression des gènes tandis que la régulation anti-sens fait intervenir un mécanisme en *trans*, c'est-à-dire agissant sur une cible hors du voisinage immédiat du riborégulateur. Les riborégulateurs liant les ARNt et le SAM sont des exemples d'un tel processus qui consiste en l'appariement, médié par la liaison du ligand, des séquences de l'aptamère du riborégulateur d'une part et d'une partie du gène régulé d'autre part (Loh et al., 2009).

### III.3. Classification

Actuellement il existe plus de 30 classes de riborégulateurs qui sont répartis en fonction de leur ligand (Serganov and Nudler, 2013). Ci-dessous sont listées les classes connues à ce jour et pour plus de facilité, elles ont été regroupées par type de ligand : acides aminés (Tab. 4), bases azotées (Tab. 5), coenzymes et cofacteurs (Tab. 6) tandis que la dernière catégorie recense ceux liant des métabolites divers (Tab. 7). Usuellement, le riborégulateur porte le nom de son ligand. La structure secondaire présentée dans les tableaux ci-dessous est celle de l'aptamère seulement, la plateforme n'étant pas une structure conservée.

Tableau 4. Riborégulateurs liant des acides aminés.

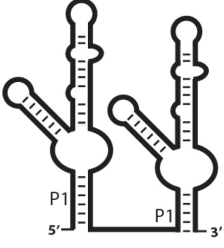
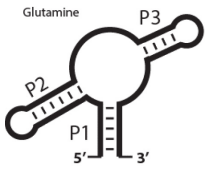
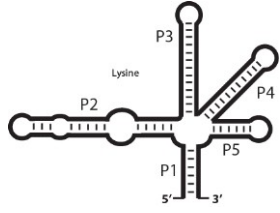
Ligand	Structure 2 <sup>nd</sup> aire	Fonctions des gènes régulés
<b>Glycine</b>  (Mandal et al., 2004)		<ul style="list-style-type: none"> <li>- Clivage de la glycine</li> <li>- Conversion et synthèse de la glycine</li> <li>- Transport</li> </ul>
<b>Glutamine</b>  (Ren et al., 2015)		<ul style="list-style-type: none"> <li>- Métabolisme de l'azote</li> </ul>
<b>Lysine</b>  (Serganov and Patel, 2009)		<ul style="list-style-type: none"> <li>- Biosynthèse</li> <li>- Transport</li> <li>- Catabolisme</li> </ul> } de la lysine

Tableau 5. Riborégulateurs liant des bases azotées.

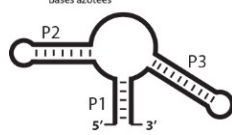
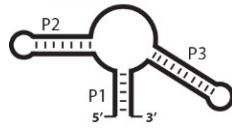
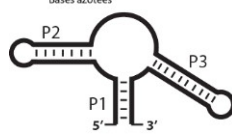
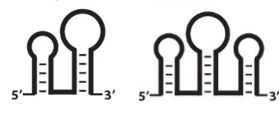
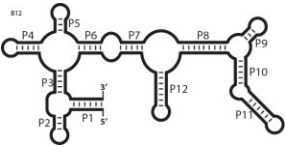
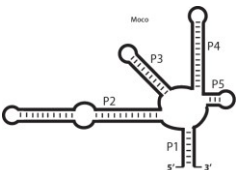
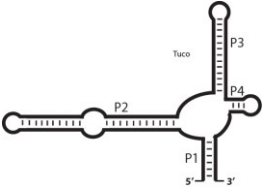
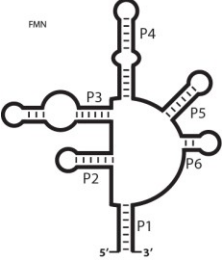
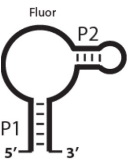
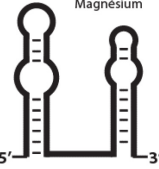
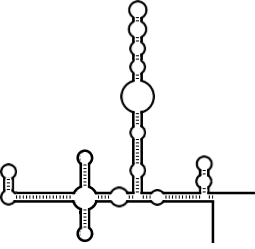
Ligand	Structure 2 <sup>nd</sup> aire	Fonctions des gènes régulés
<b>2'-déoxy-guanosine (dG)</b>  (Wacker et al., 2011)		<ul style="list-style-type: none"> <li>- Ribonucléotide réductase</li> <li>- Transporteur de phosphate</li> </ul>
<b>Adénine</b>  (Mandal and Breaker, 2004)		<ul style="list-style-type: none"> <li>- Métabolisme de l'adénine</li> <li>- Transport de protéines et de purines</li> </ul>
<b>Guanine</b>  (Kim et al., 2007; Mulhbachter et al., 2010)		<ul style="list-style-type: none"> <li>- Métabolisme</li> <li>- Transport</li> </ul> } Des purines (guanine plus particulièrement)
<b>Pré-queuosine (Pré-Q<sub>1</sub>)</b>  (Roth et al., 2007)		<ul style="list-style-type: none"> <li>- Biosynthèse de la queuosine</li> </ul>

Tableau 6. Riborégulateurs liant des coenzymes et des cofacteurs.

Ligand	Structure 2 <sup>nd</sup> aire	Fonctions des gènes régulés
<b>Adénosyl-cobalamine (AdoCbl)</b> ou <b>Coenzyme B<sub>12</sub> (CoB<sub>12</sub>)</b>  (Johnson et al., 2012)		<ul style="list-style-type: none"> <li>- Synthèse et transport de l'AdoCbl</li> <li>- Synthèse et transport de la Cbl</li> <li>- Transport du cobalt</li> <li>- Fermentation du glutamate et du succinate</li> </ul>
<b>Cofacteur de molybdène (Moco)</b>  (Regulski et al., 2008)		<ul style="list-style-type: none"> <li>- Transport du molybdate</li> </ul>
<b>Cofacteur de tungstène (Tuco)</b>  (Regulski et al., 2008)		<ul style="list-style-type: none"> <li>- Transport du tungstène</li> </ul>
<b>Flavine mononucléotide (FMN)</b>  (Pedrolli et al., 2015; Serganov et al., 2009)		<ul style="list-style-type: none"> <li>- Synthèse et transport de la riboflavine</li> </ul>
<b>Ion fluor</b>  (Baker et al., 2012)		<ul style="list-style-type: none"> <li>- Transporteur de fluor</li> <li>- Inhibition d'enzymes</li> <li>- Protection contre la toxicité du fluor</li> </ul>
<b>Ion magnésium (Mg<sup>2+</sup>)</b>  (Ramesh and Winkler, 2010)		<ul style="list-style-type: none"> <li>- Transport du Mg<sup>2+</sup></li> </ul>
<b>Ion manganèse (Mn<sup>2+</sup>)</b>  (Dambach et al., 2015; Price et al., 2015)		<ul style="list-style-type: none"> <li>- Protéines membranaires</li> <li>- Homéostasie des métaux</li> </ul>

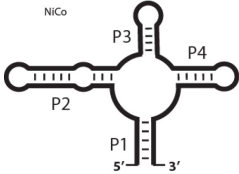
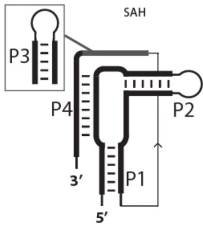
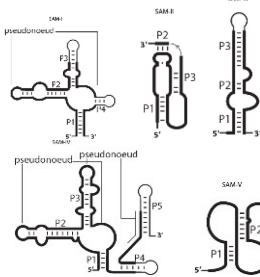
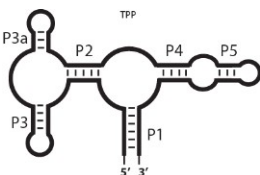
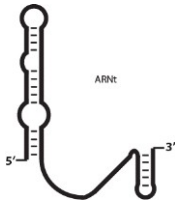
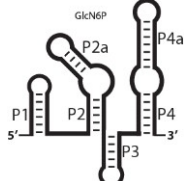
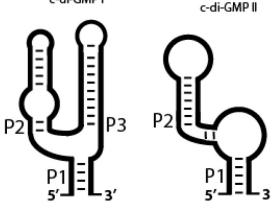
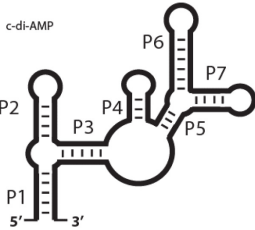
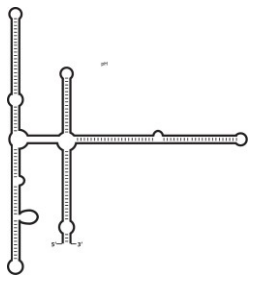

<b>Nickel et Cobalt (NiCo)</b>  (Furukawa et al., 2015)		<ul style="list-style-type: none"> <li>- Transporteurs de métaux lourds</li> <li>- Protection contre la toxicité des métaux lourds</li> </ul>
<b>S-adenosyl-homocystéine (SAH)</b>  (Edwards et al., 2010; Wang et al., 2008)		<ul style="list-style-type: none"> <li>- Conversion du SAH en homocystéine</li> </ul>
<b>S-adenosyl-méthionine (SAM)</b>  (Montange and Batey, 2006)		<ul style="list-style-type: none"> <li>- Biosynthèse de la cystéine</li> <li>- Biosynthèse et recyclage de la méthionine</li> <li>- Biosynthèse de SAM</li> <li>- Métabolisme du soufre</li> <li>- Réductase de méthylène tétrahydrofolate</li> </ul>
<b>Thiamine pyrophosphate (TPP)</b>  (Bocobza et al., 2007; Wachter et al., 2007)		<ul style="list-style-type: none"> <li>- Biosynthèse</li> <li>- Phosphorylation</li> <li>- Transport</li> </ul> } du TPP  <ul style="list-style-type: none"> <li>- Epissage d'ARN eucaryotes</li> </ul>

Tableau 7. Riborégulateurs liant des ligands divers.

Ligand	Structure 2 <sup>nd</sup> aire	Fonctions des gènes régulés
<b>ARN de transfert libre (ARNt)</b>  (Green et al., 2010)		<ul style="list-style-type: none"> <li>- Synthèse d'aminoacyl-ARNt</li> </ul>
<b>Glucosamine-6-phosphate (GlcN6P)</b>  (Collins et al., 2007)		<ul style="list-style-type: none"> <li>- Glucosamine synthase</li> </ul>
<b>Guanosine monophosphate cyclique (c-di-GMP)</b>  (Lee et al., 2010; Smith et al., 2009)		<ul style="list-style-type: none"> <li>- Environnement</li> <li>- Membranes</li> <li>- Motilité</li> </ul>

<b>Adenosine monophosphate cyclique (c-di-AMP)</b>  (Nelson et al., 2013)		<ul style="list-style-type: none"> <li>- Sporulation</li> <li>- Réponses au stress osmotique</li> <li>- Métabolisme de la paroi</li> </ul>
<b>pH</b>  (Nechooshtan et al., 2009)		<ul style="list-style-type: none"> <li>- Transporteur, protection contre des conditions alcalines extrêmes</li> </ul>
<b>Température</b>  (Johansson, 2009; Roßmanith and Narberhaus, 2016)		<ul style="list-style-type: none"> <li>- Chocs thermiques</li> <li>- Virulence</li> </ul>

Les riborégulateurs répondant à la température (ou thermosenseurs) ne sont pas reconnus unanimement comme des riborégulateurs par la communauté scientifique, certains arguant que la température peut difficilement être considérée comme un métabolite.

Outre les riborégulateurs présentés ci-dessus qui sont des riborégulateurs naturels et donc retrouvés directement dans la nature, il en existe d'autres, qui sont eux synthétiques (Berens et al., 2015; Wachsmuth et al., 2013).

#### III.4. Riborégulateurs glycine

Les riborégulateurs glycine ont été découverts en 2004 chez *B. subtilis* (Mandal et al., 2004). Ce riborégulateur permet la régulation au niveau transcriptionnel de l'opéron *gcvT* encodant pour des protéines intervenant dans le clivage de la glycine. Il peut être important pour la bactérie de réguler la concentration intracellulaire en glycine car de hautes concentrations peuvent s'avérer délétères pour la cellule. En effet, la glycine peut se substituer à l'alanine au niveau des précurseurs des peptidoglycanes voire des peptidoglycanes eux-mêmes dans les parois cellulaires. Du fait que la glycine soit un

mauvais substrat pour les enzymes en charge de la biosynthèse de la paroi, l'accumulation de précurseurs contenant de la glycine déséquilibre sévèrement la synthèse et l'hydrolyse des peptidoglycanes durant la croissance (Tezuka and Ohnishi, 2014).

Il est estimé qu'au moins 350 riborégulateurs glycine existent chez les bactéries parmi lesquels environ 60% d'entre eux régulent le clivage de la glycine via les opérons *gcvT* et *gcvP*, tandis qu'environ 20% contrôlent des transporteurs sodium/glycine ou sodium/alanine. Pour les 20% restants, leurs fonctions restent à déterminer (Ruff and Strobel, 2014). Il est néanmoins admis que les riborégulateurs permettent une réponse rapide des bactéries face à des changements de leur environnement en termes de concentration en glycine, soit en la catabolisant et en s'en servant comme source d'énergie (opéron *gcvP* chez *B. subtilis*) soit en l'important.

#### III.4.1. Structure

Le riborégulateur glycine est généralement sous forme de tandem, c'est-à-dire qu'il possède deux aptamères homologues reliés par une courte région (un *linker*) qui contrôlent une seule plateforme d'expression. Si cette architecture n'est pas commune, le riborégulateur n'est pas le seul à être en tandem (Fig. 9, A). Il existe en effet des riborégulateurs en tandem où deux riborégulateurs complets (aptamère et plateforme) se suivent comme c'est le cas pour SAM et AdoCbl chez *Bacillus clausii* (Fig. 9, B) ou encore pour deux riborégulateurs TPP chez *Bacillus anthracis* (Fig. 9, C) (Welz and Breaker, 2007). Le riborégulateur glutamine, identifié chez des cyanobactéries et des bactéries marines, peut également compter entre un et trois aptamères consécutifs. La question de savoir s'il existe qu'une seule plateforme d'expression n'est pas encore élucidée. Les séquences entre les aptamères semblent trop courtes pour comporter une plateforme d'expression mais les expériences n'ont pas encore permis de trancher la question. Les aptamères n'agissent néanmoins pas de manière coopérative (Ames and Breaker, 2011). Il semblerait donc que le riborégulateur glycine soit le seul connu à ce jour comptant deux aptamères fonctionnels régulant une seule et même plateforme d'expression.



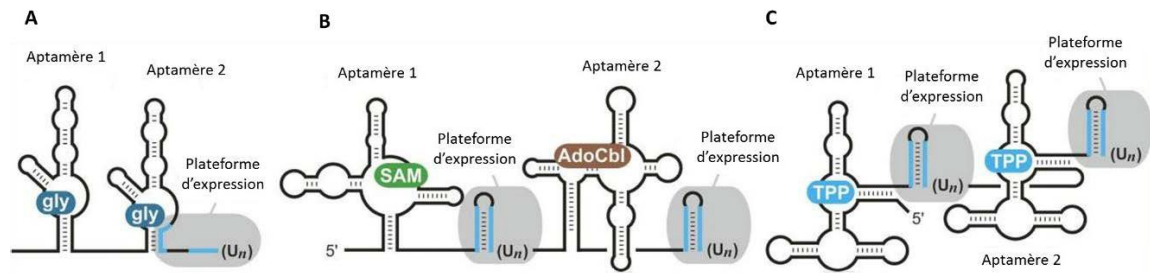


Figure 9. Exemples de riborégulateurs en tandem.

A. Riborégulateur glycine en tandem. B. Riborégulateurs SAM et AdoCbl consécutifs chez *B. clausii*.  
C. Riborégulateurs TPP consécutifs chez *B. anthracis*. Adapté de (Welz and Breaker, 2007).

Comme mentionné précédemment, les aptamères sont des structures très conservées et ont été cristallisés en 2010 chez *Vibrio cholerae* (Huang et al., 2010). On peut observer sur la Fig. 10 que certains nucléotides sont retrouvés à travers les espèces bactériennes. En effet, les nucléotides représentés en rouge sont conservés dans plus de 95% des séquences tandis que les nucléotides en bleu le sont dans plus de 75% des séquences.

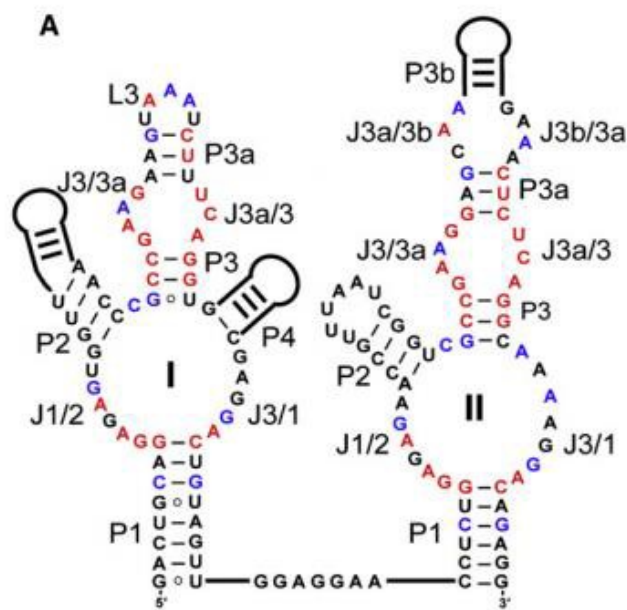


Figure 10. Structure secondaire des aptamères du riborégulateur glycine chez *V. cholerae* (Huang et al., 2010).  
Les nucléotides en rouges sont conservés dans  $\geq 95\%$  des cas tandis que ceux en bleu le sont dans  $\geq 75\%$ .

### III.4.2. Coopérativité

Initialement, il était pensé que les aptamères liaient la glycine de manière coopérative. En effet, dans des expériences de dialyse à l'équilibre avec de hautes concentrations en glycine, le second aptamère (VCII sur la Fig. 11) ou les deux (VCI-II) montrèrent des courbes de liaison de ligand sigmoïdales associées à des coefficients de Hill supérieurs à 1 ( $n=1,64$  pour *V. cholerae*, et 1,4 pour *B. subtilis* et *Fusobacterium nucleatum*). Ceci mena à l'idée que les deux aptamères fixaient la glycine de manière coopérative et positive car la valeur maximale pour deux sites de liaison est 2 (Kwon and Strobel, 2008; Mandal et al., 2004). Ce mécanisme impliquait que bien que chaque aptamère soit en mesure de lier une molécule de glycine, la liaison du premier ligand permettait un changement conformationnel à un site qui influait sur l'affinité du second aptamère pour la glycine. Un tel système permet à la cellule de réagir à de bien plus faibles variations de concentration en ligand. Des données de mutagenèse suggéraient que le sillon mineur de la P1 interagit avec le sillon majeur de la P3a afin de permettre des interactions tertiaires entre les deux aptamères (Fig. 10).

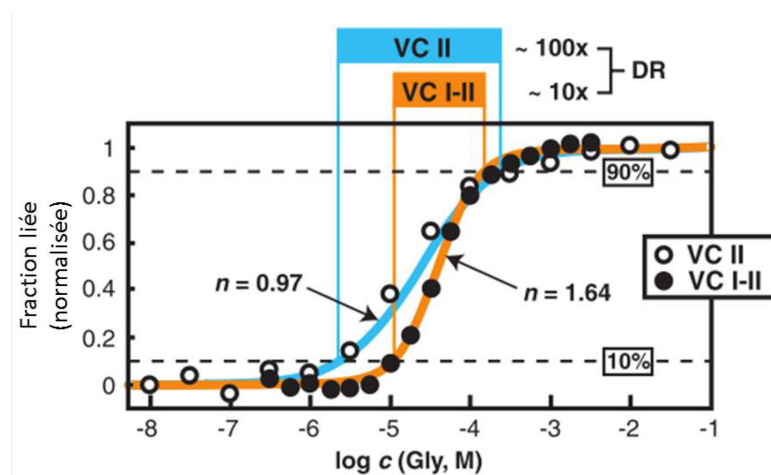


Figure 11. Liaison coopérative des deux aptamères du riborégulateur glycine chez *V. cholerae* (VC).  
Adapté de (Mandal et al., 2004)

Il a été démontré depuis que le riborégulateur glycine n'agissait pas en coopération pour la liaison de son ligand (Sherman et al., 2012). En effet, il s'agirait plutôt d'une dimérisation des aptamères.

### III.4.3. Liaison du ligand et repliement

Chaque aptamère du riborégulateur est donc en mesure de lier une molécule de glycine avec une affinité des sites de liaison de l'ordre de 1 à 30  $\mu\text{M}$ . Puisqu'il n'est plus question de coopérativité mais de dimérisation, comment celle-ci s'opère-t-elle ?

C'est la découverte d'une hélice supplémentaire en 5' (P0, Fig. 12 A), en amont des aptamères, qui a permis de constater par calorimétrie une augmentation de la liaison du ligand et d'éliminer l'idée d'une coopérativité. La liaison du ligand s'effectue au niveau de la première boucle P3 de chaque aptamère (Fig.12 B où la glycine est représentée par un ovale brun) où les nucléotides sont bien conservés d'une espèce à l'autre (Fig. 10).

Les nucléotides impliqués dans le repliement des aptamères ont été identifiés par mutagenèse d'abord chez *F. nucleatum* (Butler et al., 2011). Il existe des interactions mineures entre les aptamères entre la P1 d'un aptamère et la P3 de l'autre (a et p sur la Fig. 12 B) ce qui permet la formation d'une interface entre les deux sites de liaison du ligand (Fig.12 C). Étant donné cette dimérisation, il est attendu que les deux aptamères subissent des modifications conformationnelles avec des interactions 8 au sein même des aptamères.

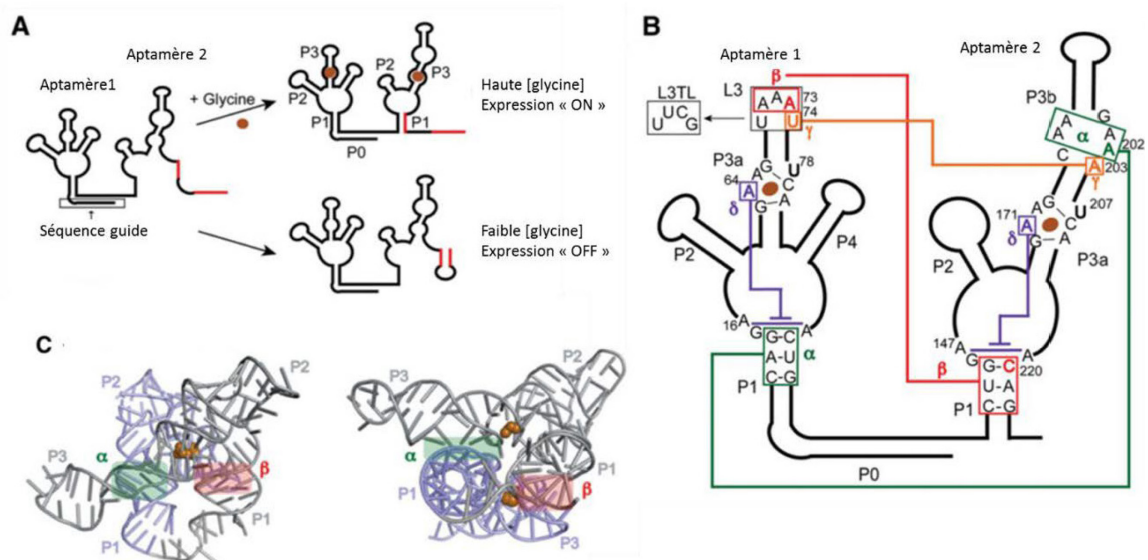


Figure 12. Structures secondaires et tertiaires du riborégulateur glycine. Adapté de (Ruff and Strobel, 2014).

#### III.4.4. Aptamères fantômes

Bien que le riborégulateur glycine soit retrouvé en tandem dans l'immense majorité des cas, il a été montré récemment que des riborégulateurs simples (appelés singlets) étaient également possibles et demeuraient fonctionnels (Ruff et al., 2016). Il s'agit possiblement de vestiges d'aptamères autrefois fonctionnels, d'où le nom de fantôme (*ghost aptamer* en anglais), qui auraient muté au fil du temps. Retrouvés en 5' ou en 3' de l'aptamère fonctionnel, ces tiges-boucles plus ou moins longues auraient encore des fonctions dans le repliement du riborégulateur, une interaction mineure étant toujours présente (Fig. 13).

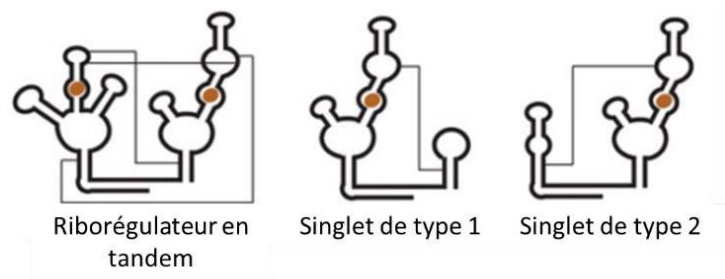


Figure 13. Architecture des riborégulateurs glycine : en tandem (gauche), en singlet (type 1 centre, type 2 à droite).  
Adapté de (Ruff et al., 2016)

## IV. *Clostridium difficile*

### IV.1. Organisme

Identifiée pour la première fois par Hall et O'Toole en 1935 dans la flore intestinale de nouveau-nés, cette bactérie s'appelait *Bacillus difficilis* en raison des difficultés rencontrées pour l'isoler et de sa croissance lente (Hall and O'Toole, 1935). Le nom de *Clostridium difficile* n'est quant à lui retrouvé qu'à partir de 1959 (McBee, 1960). Bien qu'initialement apparentée à la famille des *Clostridium*, une proposition pour renommer cette bactérie en *Peptoclostridium difficile* suite à une réorganisation de la phylogénie en 2013 a été émise (Yutin and Galperin, 2013) puis abandonnée. Elle a été depuis renommée *Clostridoides difficile* en 2016 (Lawson et al., 2016). Néanmoins, elle est toujours communément appelée *Clostridium difficile*. C'est également la nomenclature qui sera utilisée dans le présent mémoire.

Son génome a été entièrement séquencé et publié pour la première fois en 2006 par le Sanger Institute au Royaume-Uni (Sebahia et al., 2006). Il s'agissait du génome de la souche 630, une souche virulente et résistante aux antibiotiques isolée à Zurich en Suisse en 1982. Il comporte un chromosome circulaire de 4,3Mpb et un plasmide circulaire de 7881pb. Les génomes de différentes souches de *C. difficile* ont été depuis séquencés (Didelot et al., 2012; He et al., 2010). Son pourcentage en G+C est considéré faible et est de l'ordre de 29%.

Ce bacille Gram-positif anaérobie strict, pourvu de flagelles péritriches, est retrouvé de façon ubiquitaire dans la nature du fait de ses spores hautement résistantes. En effet, il a été montré que les spores peuvent survivre dans l'environnement plus de 5 mois (Kramer et al., 2006). Bien qu'associé généralement à des infections humaines, ce pathogène est en mesure d'infecter un vaste panel d'hôtes dont le bétail (cochons, chèvres, etc.) et les oiseaux d'élevage (poulets, autruches, etc.) (Moono et al., 2016). De plus, les animaux sauvages retrouvés près des élevages (rongeurs, pigeons, etc.) ont de plus hauts risques d'être infectés (Andrés-Lasheras et al., 2016).

## IV.2. Problématiques de santé publique

### IV.2.1. Historique et chiffres

De nos jours, si cette bactérie est un enjeu pour les professionnels de santé, c'est d'une part car c'est un pathogène opportuniste et d'autre part en raison de sa production de toxines. En effet, *C. difficile* possède un locus de pathogénicité (PaLoc) codant pour deux toxines : TcdA, une entérotoxine et TcdB, une cytotoxine. Néanmoins, la colite pseudomembraneuse, qui est la forme la plus sévère de l'infection, n'est pas récente.

En effet, les premiers cas identifiés de colites pseudo-membraneuses ont été rapportés dès 1893 par Finnley toutefois son occurrence était faible en raison du peu d'antibiothérapies disponibles et donc prescrites (Nolan et al., 1987). Dans les années 50, les antibiotiques commencèrent à être davantage utilisés et les cas de colites pseudo-membraneuses augmentèrent. À l'époque l'agent causatif n'était pas identifié mais *Staphylococcus aureus* était un candidat car retrouvé dans les fèces et déjà connu comme pathogène nosocomial (Chyung and Shin, 2013). Ce n'est qu'en 1978 que le lien entre des *Clostridia* produisant des toxines et la colite pseudo-membraneuse fut établi (Bartlett et al., 1978), puis le vrai coupable *C. difficile* fut identifié peu après (George, 1980; George et al., 1978).

Présent chez 5 à 15% des adultes sains, c'est l'agent principal des diarrhées survenant après antibiothérapie, rencontrées majoritairement chez les personnes âgées et/ou immunodéprimées (Galdys et al., 2014; Kobayashi et al., 2001; Ozaki et al., 2004). C'est de plus un pathogène observé fréquemment dans les hôpitaux et sur les 4,5 à 6,9% de mortalité attribuables à cette bactérie (Kwon et al., 2015; Mitchell and Gardner, 2012), 60% résultent d'infections nosocomiales. En période épidémique, cette mortalité augmente et peut atteindre 16,7% (Kwon et al., 2015). Elle est également à mettre en relation avec les nombreux cas de rechutes observés (15 à 35% pour une première rechute dont 40 à 60% feront une seconde rechute). Ces récurrences de l'infection sont imputables soit à la persistance des spores dans l'organisme soit à une nouvelle infection via l'environnement direct du patient (Cornely et al., 2012; Marsh et al., 2012). Ces rechutes résultent dans 83 à 86% des cas de la souche impliquée dans la primo-infection (Figuerola et al., 2012).

Bien que les infections à *C. difficile* (ICD) restent mortelles dans certains cas, on peut observer via les rapports de l'Institut National de Santé Publique du Québec (INSPQ) que le nombre de cas de diarrhées à *C. difficile* (DACD) est en diminution.

Tableau 8. Évolution du nombre de décès, de la létalité et des autres complications parmi les DACD nosocomiales (cat. 1a + 1b), Québec, 2011-2012 à 2015-2016 [N (%)] (données INSPQ 2016).

	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016
<b>Nombre et proportion de DACD avec suivi à 30 jours</b>	3 550 (94,0)	3 483 (91,8)	3 532 (95,7)	3 150 (91,1)	2 728 (91,6)
<b>Nombre d'installations ayant déclaré des complications</b>	91	88	90	87	84
<b>Décès</b>					
<b>Décès à 10 jours</b>	305 (8,6)	304 (8,7)	322 (9,1)	288 (9,1)	219 (8,0)
<b>Décès à 30 jours*</b>	545 (15,4)	561 (16,1)	534 (15,1)	547 (17,4)	381 (14,0)
<b>Autres complications</b>					
<b>Colectomies</b>	53 (1,5)	33 (0,9)	46 (1,3)	37 (1,2)	23 (0,8)
<b>Réadmissions</b>	155 (4,4)	182 (5,2)	206 (5,8)	205 (6,5)	185 (6,8)
<b>Transferts aux soins intensifs</b>	98 (2,8)	97 (2,8)	110 (3,1)	99 (3,1)	80 (2,9)

\*Les décès à 30 jours prennent en compte les décès à 10 jours.

D'après le *Centers for Disease Control and Prevention* (CDC, [www.cdc.gov](http://www.cdc.gov)) américain en 2015, près d'un demi-million d'ICD ont été rapportées aux États-Unis sur l'année. Parmi elles, environ 15000 morts sont directement imputables au pathogène ce qui place *C. difficile* en haut de la liste des menaces urgentes. Le CDC estime également que les dépenses de santé liées aux ICD s'élèvent à 3,8 milliards de dollars par année.

#### IV.2.1. Infection et toxines

Puisque cette bactérie est un anaérobe strict, la voie de contamination préférentielle est oro-fécale via ses spores. En effet, celles-ci vont être ingérées par le patient qu'elles soient transmises par d'autres patients (Blixt et al., 2017), les mains du personnel soignant (Jullian-Desayes et al., 2017) ou encore l'environnement (Gerba et al., 2016; Rashid et al., 2017) (Tableau 9). Là où les autres bactéries, à des stades végétatifs notamment, sont tuées par l'acidité gastrique, les spores de *C. difficile* sont suffisamment résistantes pour y survivre et peuvent ensuite germer dans l'intestin via l'action de sels biliaires qui se lient au récepteur de germination CspC (Francis et al., 2013).

Tableau 9. Répartition des cas de DACD selon l'origine présumée de l'acquisition, Québec, 2015-2016 (N, %) (données INSPQ 2016)

Catégorie	Origine de l'acquisition	N	%
<b>1a</b>	Cas relié à l'hospitalisation actuelle dans l'installation déclarante	1 998	44,1
<b>1b</b>	Cas relié à l'hospitalisation antérieure dans l'installation déclarante	979	21,6
<b>1c</b>	Cas relié aux soins ambulatoires de l'installation déclarante	230	5,1
<b>1d</b>	Cas relié à l'unité de soins de longue durée de l'installation déclarante	9	0,2
<b>2</b>	Cas relié à un séjour dans une installation non déclarante	208	4,6
<b>3</b>	Cas d'origine communautaire, non relié aux milieux de soins	909	20,1
<b>4</b>	Cas d'origine inconnue	196	4,3
	Total	4 529	100,0

S'il y a perturbation du microbiote intestinal du patient – on parle alors de dysbiose intestinale –, notamment par une antibiothérapie, une niche est alors disponible pour ce pathogène opportuniste au sein des intestins. La bactérie va alors pouvoir coloniser le milieu. Mais cette libération n'est pas le seul impact d'une dysbiose du microbiote. En effet, un microbiote sain permet la transformation des sels biliaires primaires en sels biliaires secondaires (7 $\alpha$ -déhydroxylation des dérivés de cholates en déoxycholates) qui vont inhiber la croissance de la bactérie (Fig. 14). Celle-ci ne sera alors ni en mesure de proliférer ni capable de produire ses toxines. Par contre, s'il n'y a plus cette transformation en sels biliaires secondaires (*i.e.* dysbiose), plus rien ne s'oppose à la prolifération du pathogène qui va alors pouvoir infecter son hôte (Taur and Pamer, 2014).

Après que les bactéries aient adhéré à la surface de l'épithélium, la production des toxines commence. Les deux toxines TcdA (308kDa) et TcdB (270kDa), évoquées précédemment, ont des actions cytotoxiques et pro-inflammatoires (Voth and Ballard, 2005). Ces propriétés mènent à la déstabilisation du cytosquelette d'actine et à la rupture des jonctions serrées des cellules épithéliales intestinales entraînant ainsi une perte de l'étanchéité de la barrière intestinale. Plus précisément, les toxines sont endocytées après leur libération par la bactérie, puis sont ensuite transloquées dans le cytosol et vont alors inactiver de petites protéines liant le GTP dont Rho, Rac et Cdc42. Cette inactivation via monoglucosylation d'un résidu thréonine retrouvé dans la boucle de liaison à l'effecteur et coordonnant un cation divalent, est critique pour la liaison du GTP. En glucosylant de petites GTPases, ces toxines entraînent la condensation de l'actine et l'arrondissement des cellules, causant alors leur mort. En plus de l'apoptose des cellules, ceci conduit à



une inflammation des cellules épithéliales de l'intestin à laquelle le système immunitaire de l'hôte répond via un panel d'agents tels des anticorps et des cytokines. Cette réponse résulte en une augmentation de la sécrétion de fluides à travers la barrière intestinale, d'où les diarrhées.

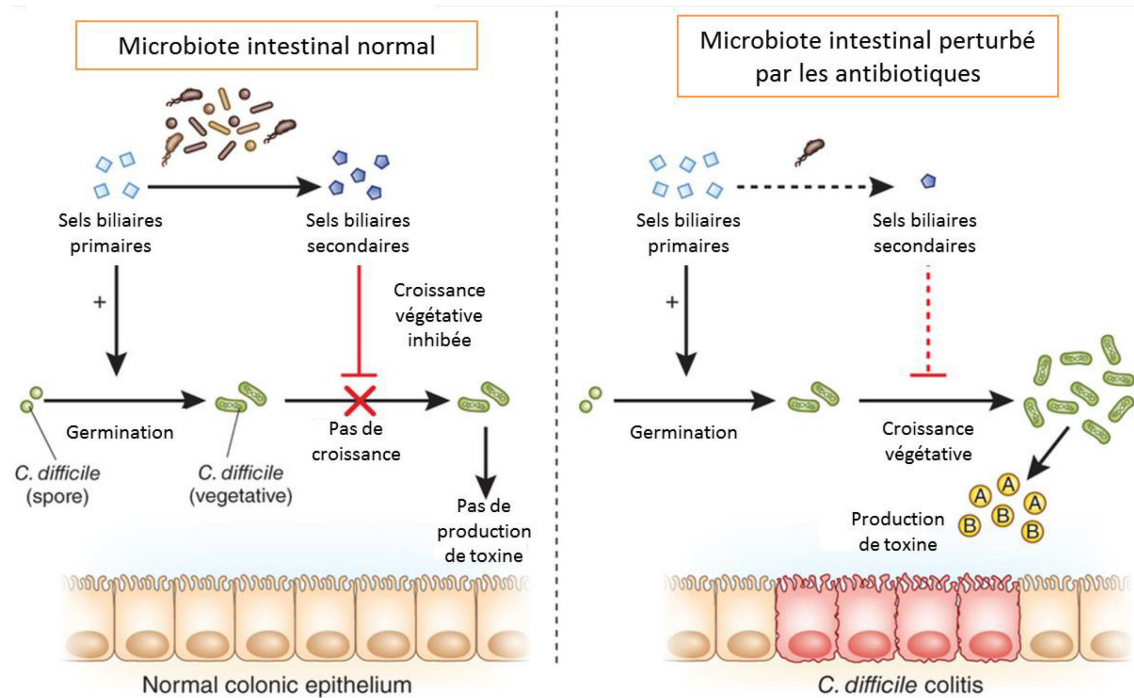


Figure 14. Germination et production de toxines par *C. difficile* en présence d'un microbiote sain ou perturbé. (Adapté de (Taur and Pamer, 2014))

Mais le système immunitaire ne doit pas seulement lutter contre cette inflammation. En effet, l'accumulation de cellules mortes le long du colon forme une pseudomembrane qui va bloquer le passage des anticorps, entre autres, du sang vers les intestins.

Un faible pourcentage de *C. difficile* va également sporuler dans les intestins puis être libéré dans les fèces pour potentiellement contaminer d'autres patients via une autre contamination oro-fécale.

#### IV.2.2. Pathologies et traitements

##### IV.2.2.1. Pathologies

Les ICD se traduisent par un panel de symptômes dont l'évolution la plus sévère et parfois mortelle est la colite pseudo-membraneuse. Mais on note également l'apparition de fièvres, de nausées, de douleurs abdominales ainsi que de diarrhées aqueuses parfois sanglantes et d'une leucocytose. De plus il peut survenir en cas d'infection sévère une complication qu'on appelle le mégacôlon toxique s'accompagnant d'un risque de perforation des intestins. Le mégacôlon toxique est caractérisé par un côlon très dilaté et des gonflements abdominaux en sus de la fièvre et des douleurs abdominales.

Les médecins ont à leur disposition un certain nombre d'examens qu'ils peuvent utiliser pour diagnostiquer l'infection (Hall and Premji, 2015). Pour détecter la présence de la bactérie, il est possible de procéder à une réaction de polymérase en chaîne (PCR) directement sur les fèces. Sur ces mêmes fèces, un essai immunologique peut être fait pour détecter les toxines tandis que des radiographies de l'abdomen peuvent permettre de déceler une dilatation du colon.

##### IV.2.2.2. Traitements

Les traitements les plus couramment prescrits en cas d'ICD sont à base de métronidazole ou de vancomycine, deux antibiotiques à large spectre (Spigaglia, 2016). Surtout depuis que les souches épidémiques (027/B1/NAP1) sont devenues résistantes aux fluoroquinolones, des antibiotiques qui ciblent l'ADN gyrase bactérienne. Ce sont ces résistances qui sont à l'origine des épidémies meurtrières au début des années 2000 à travers le monde (He et al., 2013) (Fig. 15).

La vancomycine est un antibiotique bactéricide de la famille des glycopeptides et agit en inhibant la synthèse du peptidoglycane de la paroi bactérienne. En raison de sa taille (1500 à 2000 daltons), cette molécule n'est pas en mesure d'emprunter la voie des porines des bactéries Gram-négatives et possède donc un spectre d'action restreint aux Gram-positives.

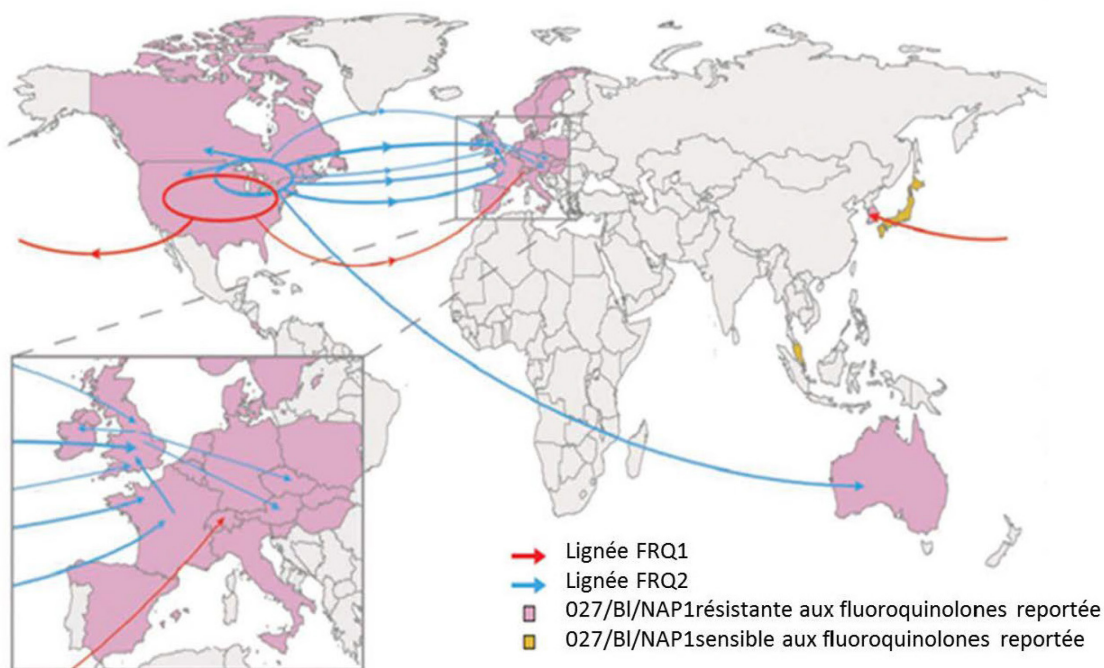


Figure 15. Carte montrant la dispersion des deux lignées épidémiques de *C. difficile* résistantes aux fluoroquinolones à travers le monde (He et al., 2013).

Le métronidazole est, quant à lui, un antibiotique mais aussi un antiparasitaire de la famille des nitroimidazoles. Il fonctionne en inhibant la synthèse des acides nucléiques. Son spectre d'action englobe les bactéries anaérobies et certains protozoaires.

La fidaxomycine est une molécule récente (Vaishnavi, 2015) et présente un spectre d'action étroit puisqu'elle n'éradique que les Gram-positives et plus particulièrement les Clostridies avec un minimum de perturbations sur le microbiote sain. Elle agit en se fixant et en prévenant les mouvements des régions « *switch* » de l'ARN polymérase bactérienne. Ces régions permettent d'ouvrir et de fermer les interactions ARN/ADN notamment lors de l'initiation de la transcription. La fidaxomycine a pour avantage de diminuer les risques de rechutes, ce qui compense en partie son coût plus élevé comparativement à la vancomycine et au métronidazole (Watt et al., 2017).

Le traitement prescrit va toutefois dépendre de la sévérité de l'infection et est décrit ci-dessous :

Tableau 10. Traitement conseillé en fonction de la sévérité de l'infection

Caractéristique de l'infection	Thérapie recommandée
<b>Infection bénigne</b> (Pas de symptômes systémiques, diarrhée bénigne)	Métronidazole 250 mg oralement 4x/j Ou 500 mg 3x/j pendant 10j
<b>Infection modérée</b> (Fièvre, diarrhée abondante, douleur abdominale, leucocytose)	Vancomycine 125-500 mg oralement 4x/j pendant 10j
<b>Infection sévère</b> (Iléus paralytique, mégacolon toxique, déshydratation ou sepsis)	Consultation chirurgicale + vancomycine intraluminale
<b>Incapacité à prendre une médication orale</b>	Vancomycine intraluminale avec ou sans métronidazole par IV

Tableau 11. Coût des antibiotiques en fonction de leur mode d'administration

	Coût par dose	Régime	Coût pour 10j
Métronidazole 500 mg	\$0,73	500 mg 3x/j	\$22,00
Vancomycine 125 mg (cachets)	\$17,00	125 mg 4x/j	\$680,00
Vancomycine 125 mg IV	\$2,50-\$10,00	125 mg 4x/j	\$100,00-400,00
Fidaxomyline 200 mg	\$140,00	200 mg 2x/j	\$2800,00

Si le traitement au métronidazole reste peu coûteux, la vancomycine et la fidaxomyline sont quant à elles bien plus chères et représentent des charges importantes pour les systèmes de santé. Devant l'apparition de plus en plus importantes de résistances chez la bactérie, d'autres thérapies ont vu le jour ou ont été redécouvertes parmi lesquelles peuvent être citées la transplantation fécale ou la thérapie par les phages.

#### IV.2.2.2.1. La transplantation fécale

La transplantation fécale, aussi appelée bactériothérapie fécale, a été utilisée pour la première fois en 1958 par l'équipe de Ben Eiseman pour le traitement de la colite pseudo-membraneuse. Elle consiste à administrer des matières fécales homogénéisées dans de l'eau puis filtrées d'un individu sain vers le tractus gastro-intestinal d'un individu malade via des énémas (Eiseman et al., 1958). Depuis, d'autres méthodes ont été développées telles des tubes nasogastriques ou nasojejunaux, des énémas de rétention fécale ou encore la colonoscopie. Elle permet de réintroduire une communauté microbienne stable et saine pour remplacer ou aider un microbiote intestinal défaillant. En effet, une ICD se traduit souvent par une diminution de la diversité bactérienne. Cette

approche s'est révélée être très efficace contre les ICD sévères et/ou récurrentes (Costello et al., 2016; Edmond, 2016).

Pourtant, ce n'est pas une technique récente. En effet, on retrace ses premières utilisations au IV<sup>ème</sup> siècle en Chine par le médecin Ge Hong (relaté dans le premier écrit de médecine d'urgence le « Zhou Hou Bei Ji Fang ») où elle était utilisée avec succès chez l'humain pour soigner des diarrhées sévères ou des intoxications alimentaires. Ce traitement, composé de matières fécales et d'eau, devait être bu par le patient (Zhang et al., 2012). Au 16<sup>ème</sup> siècle cette méthode est toujours employée et a vu ses applications élargies aux pathologies abdominales impliquant fièvre, vomissements, fièvres et constipations. De plus, au lieu de se limiter aux fèces fraîches, elles furent également utilisées fermentées ou séchées. Néanmoins Li Zheng, dont les écrits nous sont parvenus, a préféré renommer la mixture « soupe dorée » pour des raisons esthétiques et certainement pour obtenir l'accord de ses patients pour l'administrer. Plus récemment, les fèces chaudes de chameaux étaient consommées par les Bédouins en cas de dysenterie bactérienne, ce que confirment quelques soldats allemands des troupes africaines durant la seconde guerre mondiale (Lewin, 2001).

Bien que toujours employée en médecine chinoise, elle n'était utilisée qu'en médecine vétérinaire en Occident et ce, en prévention de la salmonellose chez les poulets notamment (Nurmi and Rantala, 1973). Ce n'est que dans les années 50 qu'un regain d'intérêt eut lieu.

De nos jours, la transplantation fécale n'est que peu prescrite chez l'humain et seulement pour des ICD intraitables par les antibiotiques actuels. Pourtant de très bons résultats sont obtenus et le taux de guérison dépasse les 90% (Edmond, 2016). Cette utilisation marginale peut s'expliquer en partie par le fait qu'elle n'est reconnue en tant que traitement aux États-Unis que depuis 2013. De plus, les donneurs étaient préférentiellement des amis ou la famille du patient ce qui pouvait s'avérer problématique. Les patients pouvaient être gênés ou ne disposaient pas de proches à qui demander ce service. De plus certains donneurs pouvaient être rétifs à accepter si pour des raisons médicales diverses ils devaient être refusés comme par exemple être HIV-positif (Costello et al., 2016). Depuis quelques années, une banque conservant des échantillons congelés de fèces a été créée ([openbiome.org](http://openbiome.org)) et permet une anonymisation des donneurs, une standardisation des dons et la distribution à une plus grande échelle.

Le nombre de donneurs reste toutefois restreint et ne permet pas d'envisager ce traitement de manière courante. En effet, une batterie de tests est conduite sur les échantillons pour s'assurer qu'ils ne seront pas vecteurs de maladies ou de virus (Hépatite, VIH, porteur sain pour un pathogène, etc.) (Woodworth et al., 2017). Les critères de sélection sont nombreux mais permettent d'assurer un traitement optimal au patient. Néanmoins quelques problèmes sont survenus depuis que cette technique est employée. Ils sont certes très rares voire anecdotiques mais au moins un cas d'obésité survenu après la transplantation provenant d'un donneur obèse mais en bonne santé a été constaté (Alang and Kelly, 2015). Des conditions plus graves ont également été observées comme un syndrome septique passager (Singh et al., 2016).

#### IV.2.2.2.2. La phagothérapie ou thérapie par les phages

Les phages, virus infectant uniquement les bactéries, ont été découverts en 1909 et 1915 par Félix d'Hérelle, un microbiologiste français et Frederick Twort, un microbiologiste britannique respectivement (Chanishvili, 2012). Durant ses années au Mexique (1906-1909), d'Hérelle avait observé des plaques de lyse dans des cultures d'*Enterobacter aerogenes* isolées d'épidémies chez les locustes. Il laissa ces plaques qu'il qualifiait de « tâches vierges » de côté sans comprendre de quoi il s'agissait. Lorsque les travaux de Twort furent publiés en 1915 sur des lyses bactériennes transmissibles (Twort, 1915), d'Hérelle fut en mesure de comprendre que ces tâches vierges résultaient en fait de virus bactériens. En effet, Twort décrivait dans son article des colonies de *Staphylococcus* dont les bords étaient grignotés. Il avait eu l'idée de filtrer les cultures et de transférer le surnageant sur des cultures fraîches de *Staphylococcus* où il constata également des plaques de lyse.

La thérapie par les phages utilise la spécificité des phages pour détruire une bactérie donnée. La spécificité des phages est tellement grande qu'elle peut ne pas s'étendre à toutes les souches d'une même espèce. L'utilisation de phages permet donc de cibler une bactérie au sein d'un microbiote sans impacter sur les autres bactéries. Néanmoins, des cocktails de phages sont utilisés afin d'augmenter le spectre d'action d'un tel traitement à une espèce donnée.

Si en Occident cette thérapie est restée peu étudiée et employée, surtout depuis l'ère des antibiotiques débutée dans les années 40 (quelques études sont parues auparavant sur

la méningite par exemple (Eaton and Bayne-Jones, 1934; Schless et al., 1932)), ce n'est pas le cas dans l'ancienne URSS où la phagothérapie a été largement utilisée notamment lors de la seconde guerre mondiale avec de très bons résultats (Chanishvili, 2012). Ainsi, le traitement de la dysenterie, de la fièvre typhoïde, des gangrènes gazeuses et des blessures par des cocktails de phages que ce soit administrés oralement ou appliqués localement sur la blessure ont permis d'améliorer grandement la survie des patients et de diminuer la nécessité des amputations. De même les épidémies de choléra en Afghanistan dans les années 58-60 ont pu être largement contenues via la prophylaxie (Sayamov, 1963). Néanmoins, presque toute la littérature s'y rapportant étant rédigée en russe et le peu de communications durant la guerre froide limitaient grandement les échanges entre Occident et le bloc communiste et ces avancées sont restées longtemps méconnues.

Pour qu'une thérapie par les phages soit acceptée, il faut normalement que les phages utilisés soient de nature lytique, c'est-à-dire qu'ils doivent nécessairement tuer la bactérie qu'ils infectent. Bien que des essais aient été entrepris avec succès chez *C. difficile* (Meader et al., 2013; Nale et al., 2016), le problème principal est l'absence de tels phages. L'utilisation de cocktails de phages combinant des myovirus et des siphovirus a permis la lyse complète de *C. difficile in vitro* sans apparition de colonies résistantes. Des essais chez un modèle de hamster ont également permis de diminuer la colonisation de la bactérie après 36h. Il y a donc un réel potentiel dans l'utilisation des phages pour lutter contre *C. difficile*. De plus des travaux dans notre laboratoire (Fortier) ont été entrepris afin de muter l'intégrase pour empêcher l'intégration du génome viral au sein de l'ADN bactérien et le répresseur cI de phages lysogéniques afin d'en faire des phages lytiques.

#### IV.3. *Clostridium difficile* souche 630 et riborégulateurs

Comme évoqué précédemment (IV.1.), la souche 630 de *C. difficile* (NC\_009089.1) est une souche virulente responsable d'infections chez l'humain. Néanmoins, d'autres souches virulentes et épidémiques existent et si celle-ci a été choisie comme modèle d'étude au sein du laboratoire c'est en raison des outils moléculaires disponibles pour manipuler cette souche ainsi que d'une étude transcriptomique parue récemment et servant de base de travail (Soutourina et al., 2013).

Sur les 66 riborégulateurs identifiés par bio-informatique chez *C. difficile*, seuls deux sont en mesure de lier la glycine : celui en amont du symporteur sodium/glycine (GeneID : 4916417) et la glycine réductase *grdX* (GeneID: 4915151). Les riborégulateurs demeurent encore peu étudiés chez cette bactérie, les études portant davantage sur des organismes modèles comme *E. coli* ou *B. subtilis*.



## V. Objectifs de maîtrise

Les stratégies actuelles employées pour lutter contre *C. difficile* s'avèrent insuffisantes, il est donc nécessaire d'apporter de nouvelles pistes. Une approche intéressante pour le développement de nouveaux antibiotiques est l'utilisation des riborégulateurs. En effet, si le blocage ou la surexpression du gène contrôlé par le riborégulateur s'avère néfaste pour la bactérie et impacte sa survie et/ou sa compétitivité et donc son pouvoir pathogène, le développement de molécules analogues non-métabolisables du ligand, ici la glycine, pourrait être utilisé à des fins thérapeutiques. Cette stratégie a déjà été utilisée avec succès pour la mammite bovine causée par *S. aureus* via un analogue de la guanine (Ster et al., 2013) mais aussi récemment chez la souris pour lutter contre des infections à *C. difficile* grâce à des analogues liant le riborégulateur FMN (Blount et al., 2015).

Les objectifs de cette maîtrise se divisent en deux parties. D'une part étudier le fonctionnement de ces riborégulateurs glycine atypiques et de les caractériser *in vitro* afin de valider les données bio-informatiques servant de bases à ce projet via des techniques telles la cartographie chimique ou la transcription à cycle unique. Le fonctionnement des riborégulateurs est également étudié *in vivo* à l'aide d'essais rapporteurs. D'autre part il s'agit également de déterminer si ceux-ci constituent des cibles intéressantes pour une potentielle approche thérapeutique en réalisant des mutants des gènes contrôlés chez *C. difficile* pour évaluer leur importance chez la bactérie.

## VI. Matériels & Méthodes

### VI.1. Souches utilisées

La souche de *C. difficile* utilisée durant les expériences est la souche 630 $\Delta$ *erm* (NZ\_LN614756.1) (Sebaihia et al., 2006) fournie par le laboratoire de Nigel Minton (Nottingham, UK). *C. difficile* possède deux gènes de résistance pour l'érythromycine (*erm*) et cette souche n'en a plus qu'un de fonctionnel (*ermI*) ce qui est toutefois insuffisant pour assurer une protection. Du fait de l'impossibilité de transformer directement *C. difficile*, des souches de *E. coli* ont été utilisées pour les constructions plasmidiques et clonages avant de transférer les plasmides obtenus dans *C. difficile* 630 $\Delta$ *erm* par conjugaison. Ces souches sont la CA434 (HB101), qui possède un plasmide de conjugaison (R702 possédant la machinerie conjugative) et la TOP10 (génotype : *F-mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*) $\phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74 nupG recA1 araD139*  $\Delta$ (*ara-leu*)7697 *galE15 galK16 rpsL(StrR)endA1*  $\lambda$ -) (Invitrogen, Burlington, ON) servant pour la conservation des plasmides à -80°C dans des stocks glycérolés.

Toutes les expérimentations impliquant des souches de *C. difficile* ont été réalisées au sein d'une chambre anaérobie en vinyl (Coy Laboratory Products, Inc.) remplie d'un mélange de diazote (85%) et de gaz mix (5% CO<sub>2</sub> et 10% H<sub>2</sub>). Les cultures s'effectuent à 37°C au sein d'un incubateur présent dans la chambre et sans agitation. Tous les milieux et solutions employés nécessitent d'être dégazés (minimum 5h) dans la chambre avant utilisation pour s'assurer de l'absence d'oxygène.

Dans le cas des essais  $\beta$ -galactosidase, une souche d'*E. coli* (DH5 $\alpha$ ) est également utilisée pour la construction des plasmides avant que celui-ci ne soit transféré dans la souche de *Bacillus subtilis* d'intérêt, la souche 168 (NC\_000964.3) (Zeigler et al., 2008).

### VI.2. Préparation des cellules compétentes

Les cellules d'*E. coli* compétentes (TOP10 ou CA434) sont préparées selon un protocole interne au laboratoire. À partir d'une culture en bouillon riche, 1 mL est prélevé pour inoculer 100 mL de milieu Psi (pour 100 mL : 0,5 g Extrait de levure, 2 g tryptone, 0,5 g MgSO<sub>4</sub>, pH 7,6 ajusté au KOH) puis le milieu est incubé à 37°C avec agitation jusqu'à atteindre une densité optique à 600nm (DO<sub>600</sub>) de 0,5. Les cellules sont alors

misées sur glace 15min avant d'être centrifugées à 3700g pendant 5 min. Le culot est repris dans 40 mL de milieu TbfI (60 mM  $\text{KCH}_3\text{COO}$  ; 200 mM  $\text{RbCl}$  ; 26,5 mM  $\text{CaCl}$  ; 160 mM  $\text{MnCl}$ , 30 mL de glycérol ; pH 5,8 ajusté à l'acide acétique puis filtré) puis remis sur glace pour 15 min et centrifugé comme précédemment avant d'être resuspendu dans 4 mL de TfbII (10 mM MOPS ; 100 mM  $\text{CaCl}$  ; 10 mM  $\text{RbCl}$  ; 15 mL de glycérol ; pH 6,5 ajusté au NaOH). La suspension bactérienne est alors aliquotée et congelée à  $-80^\circ\text{C}$  pour utilisation ultérieure.

Les cellules de *B. subtilis* compétentes sont préparées selon le protocole décrit par (Vojcic et al., 2012) pour la souche DB104 et adapté pour la 168. Les bactéries sont inoculées dans 10 mL de milieu SM1 (*Starvation medium* 1) puis incubées à  $37^\circ\text{C}$  sous agitation pendant la nuit (O/N). Les cultures sont ensuite diluées pour obtenir une  $\text{DO}_{600}$  de 0,5 dans un volume final de 10 mL de SM1 avant d'être incubées pendant 3h à  $37^\circ\text{C}$  sous agitation. À cette culture sont ajoutés 10 mL de SM2 supplémenté de 50  $\mu\text{g/mL}$  de tryptophane, la souche 168 étant auxotrophe pour cet acide aminé. Après 2h d'incubation dans les mêmes conditions, 4 mL de glycérol 50% sont ajoutés et la culture est congelée à  $-80^\circ\text{C}$  à raison de 500  $\mu\text{L}$  par cryotube. Le milieu SM1 contient 0,2% de sulfate d'ammonium ; 1,4%  $\text{K}_2\text{HPO}_4$  ; 0,6%  $\text{KH}_2\text{PO}_4$  ; 0,07% de citrate de sodium ; 0,5% de glucose ; 0,02%  $\text{MgSO}_4,7\text{H}_2\text{O}$  ; 0,2% d'extraits de levure et 0,025% d'acides casaminés. Le milieu SM2 quant à lui est plus pauvre en nutriment et contient seulement 0,1% d'extraits de levure, 0,01% d'acides casaminés en plus de 0,05%  $\text{NaCl}$  et de 0,08% de  $\text{MgSO}_4,7\text{H}_2\text{O}$ . Les solutions sont autoclavées mais le glucose, le  $\text{MgSO}_4,7\text{H}_2\text{O}$  et le  $\text{NaCl}$  sont ajoutés stérilisés par filtration puis ajoutés stérilement par la suite.

### VI.3. Extraction d'ADN génomique

L'ADN génomique de *C. difficile* a été extrait afin de réaliser les PCR et autres techniques nécessitant de travailler directement sur l'ADN. Pour cela, 10 mL de culture de la souche d'intérêt de *C. difficile* sontensemencés dans du BHI O/N, puis au matin sont centrifugés 10 min/3700g. Le culot est ensuite congelé 15 min à  $-80^\circ\text{C}$  afin de fragiliser la paroi avant d'être dégelé et repris dans 500  $\mu\text{L}$  de PBS 1X auxquels sont ajoutés 50  $\mu\text{g/mL}$  d'endolysine, enzyme lytique clonée d'un bactériophage et purifiée par le service de purification de protéines du département. Une phase d'incubation à  $37^\circ\text{C}$

s'ensuit pendant 30 à 60 min jusqu'à ce que la solution devienne presque limpide. Sont alors ajoutés 20 µL de protéinase K à 20 µg/µL et le tube est incubé 15 à 20 min à 55°C.

À cela sont adjoints 500 µL de phénol/chloroforme (1:1), le mélange est alors émulsifié par inversion avant d'être centrifugé 5 min/16200g. Le surnageant est récupéré sans toucher l'interface puis transféré dans un nouveau tube et l'opération est répétée 2 fois. Pour enlever toutes traces résiduelles de phénol, une dernière extraction avec seulement du chloroforme est effectué puis l'ADN est précipité via l'ajout de 1/10 de volume d'acétate de sodium 3 M et de 2 volumes d'éthanol 100% avant d'être centrifugé comme précédemment. Dans le cas d'apparition d'un précipité blanchâtre, représentant l'ADN, celui-ci est récupéré manuellement à l'aide d'une pointe de pipette ou s'il n'y a pas de précipité visible, le tube est placé à -20°C pendant 30 min. Dans ce dernier cas, le tube est alors centrifugé 10 min/16200g (ou 5 min à 3700g si la congélation a permis l'apparition du précipité car une vitesse trop importante aura tendance à faire précipiter les sels également). Le culot est alors lavé à l'éthanol 70% puis centrifugé 2 min/16200g avant de sécher à l'air libre. Il est ensuite solubilisé dans 200 µL de Tris-HCl 10 mM et traité avec 10 µL de RNase à 10 mg/mL pendant 15 min à 37°C avant d'être dosé et vérifié sur gel d'agarose 0,8%.

#### VI.4. Réaction de polymérisation en chaine (PCR)

L'amplification des gènes ou séquences d'ADN nécessaires à la réalisation des différentes expériences décrites ultérieurement s'effectue par réaction de polymérisation en chaine (*polymerase chain reaction*). Elle consiste en une phase initiale de dénaturation thermique suivie par 30 ou 35 cycles de 3 étapes (Tab. 12) : dénaturation thermique, appariement de la matrice d'ADN avec l'amorce et enfin élongation de l'amorce à l'aide d'une polymérase. Au terme de ces cycles, une étape d'extension finale est effectuée avant le refroidissement des produits à 4°C brièvement.

Tableau 12. Conditions de PCR utilisées. X différenciant selon le Tm des couples d'amorces utilisés

Étape	Température	Durée
Dénaturation initiale	94°C	3min
Dénaturation	94°C	30 sec
Appariement	X	45 sec
Élongation	72°C	1min/kb d'ADN
Extension finale	72°C	5min
	4°C	1min

Les réactions s'effectuent à partir de 10 ng d'ADN génomique (ADNg), 1X de tampon correspondant à la polymérase employée, 0,1  $\mu$ M d'amorces ainsi que 2,5 unités de polymérase d'ADN (Taq ou Pfu maison ou commerciales). Elles ont été réalisées dans des thermocycleurs Eppendorf Mastercycler ou avec un thermocycleur TGradient de Biomera. Les amorces quant à elles ont été synthétisées via les compagnies Integrated DNA Technologies (IDT) pour celles avec le code LCF tandis que les autres ont été synthétisées par Bio Basic Inc.

Les produits des amplifications ont ensuite été contrôlés sur gel d'agarose 1 ou 2% (p/v) en fonction de la taille attendue (>1,5kb et <1,5kb respectivement). Les gels sont préparés dans du tampon TAE auquel est ajouté au moment de couler du bromure d'éthidium. Ceci permet ensuite la visualisation des fragments d'ADN à l'aide d'une lampe UV (ImageQuant IQ 300 de GE Healthcare).

## VI.5. Construction de plasmides : Gibson et ligation

### VI.5.1 Gibson

La réaction Gibson, nommée ainsi d'après le premier auteur de l'article décrivant cette technique, est une réaction enzymatique permettant l'assemblage directionnel de molécules d'ADN pouvant atteindre plusieurs centaines de kilobases (Gibson et al., 2009). Dans le cadre des travaux présentés ici, elle utilise des amorces chevauchantes entre le plasmide (pRPF144E, plasmide identique au pRPF144 (Annexe 2) mais dépourvu du gène codant pour l'enzyme glucuronidase) et l'insert d'ADN, à raison de 20 nucléotides sur chaque ADN. Une PCR réalisée sur l'ADN génomique permet la synthèse

d'un insert avec des séquences complémentaires au vecteur. Le vecteur est pour sa part linéarisé à l'aide de l'enzyme BamHI dans notre cas. Les deux ADN (plasmide et insert) sont ajoutés de manière équimolaire pour 100ng de vecteur et sont ajoutés au mélange réactionnel Gibson à un ratio v/v de 1:1. La solution est ensuite incubée 1h à 50°C puis diluée 1:5 pour transformer des *E. coli* TOP10.

Le mélange réactionnel de Gibson contient plusieurs enzymes aux fonctions bien spécifiques. Tout d'abord, la T5 exonuclease va dégrader les extrémités 5' des ADN mais étant donné que l'incubation s'effectue à 50°C, sa vitesse d'action est lente (sa température optimale est 37°C). L'enzyme étant thermolabile, elle est peu à peu inactivée à 50°C. La polymérase Phusion va alors prendre le relais et synthétiser les brins complémentaires en associant l'insert et le plasmide puis la *Taq* ligase va compléter la liaison entre les segments d'ADN (Fig. 16).

Les constructions sont vérifiées par PCR et/ou digestion enzymatique car le site de restriction BamHI a été reconstitué au sein de chaque amorce.

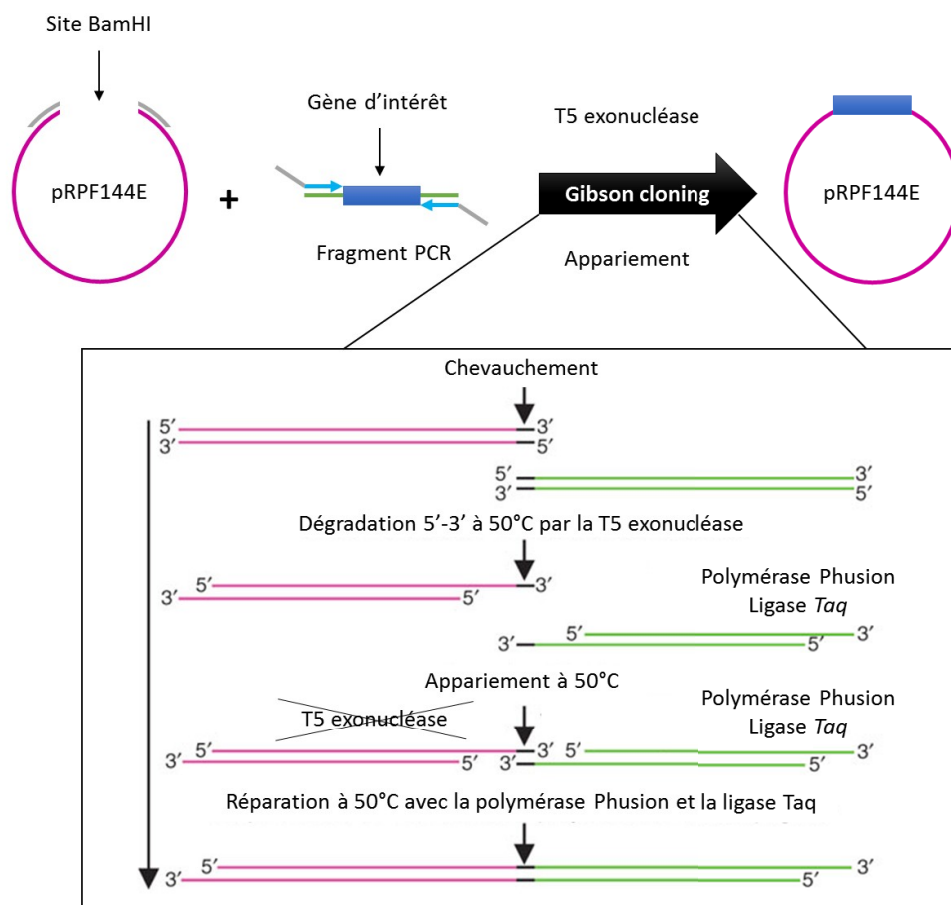


Figure 16. Principe de la réaction Gibson. Adapté de (Gibson et al., 2009).

### VI.5.2. Ligation

Les ligations ont été réalisées dans le cadre de la construction des plasmides pour les essais glucuronidase. Pour ce faire, des amorces ont été dessinées pour inclure un site de restriction précédé en 5' d'environ 6 nucléotides. Les sites de restriction sont SacI et BamHI et correspondent à ceux retrouvés en amont du gène *lacZ* sur le plasmide pDG1661 (Annexe 3). Après digestions enzymatiques des produits de PCR et du plasmide (SacI-HF et BamHI-HF, NEB), les ADN sont purifiés sur colonne. La ligation est réalisée via l'enzyme T4 DNA ligase avec 200ng de plasmide et 25ng d'insert. Les réactions sont ensuite purifiées puis validées par PCR. Les plasmides sont transférés dans les *E. coli* DH5 $\alpha$  selon le protocole donné en VI.7.1.

### VI.6. Extraction rapide de plasmide à la silice (*Miniprep*)

La silice est préparée à partir de 2g de silice (Sigma S5631) suspendue dans 15 mL d'eau distillée. Après sédimentation, le surnageant laiteux est ôté et mis à sédimenter deux autres fois. Une fois le culot bien formé, il est repris dans 50 mL d'eau ultra-pure et conservé à température pièce.

#### VI.6.1. Chez *E. coli*

Afin de vérifier les constructions plasmidiques et de les transférer dans les autres bactéries (*E. coli* CA434 et *C. difficile* 630 $\Delta$ erm), il est important de pouvoir les extraire de la souche d'*E. coli* TOP10 ayant servi à les construire. Pour cela, nous disposons d'un protocole propre au laboratoire.

À partir d'une culture O/N (LB, 37°C, 220 rpm) de *E. coli* TOP10 contenant le plasmide, 1,5 à 3 mL sont centrifugés 1 min/16200g. Au culot sont ajoutés 200  $\mu$ L de tampon P1 (50 mM Tris-HCl pH 8, 10 mM EDTA) puis mélangés par aspiration-refoulement avant que 200 $\mu$ L de tampon P2 (200 mM NaOH, 1% SDS) et 200  $\mu$ L de tampon P3 (KoAc 3M, pH 4,8) ne soient ajoutés successivement et mélangés par inversion. Le mélange est alors centrifugé 5 min/16200g et le surnageant est ajouté à 200  $\mu$ L de silice préalablement aliquotée dans un nouveau tube. Après mélange par inversion, le tube est centrifugé 10 sec/16200g et la silice est lavée avec 500  $\mu$ L de tampon

de lavage (50% éthanol, 0,1 M NaCl, 1 mM EDTA, 10 mM Tris-HCl pH 7,5) puis centrifugée 10 sec/16200g. Elle est ensuite laissée sécher à l'air libre avant que l'ADN ne soit élué via 50 µL de Tris-HCl 10 mM pH 8 et incubé 3 min à 55°C. L'ADN est alors récupéré par centrifugation 1 min/16200g et le surnageant transféré avant de subir une nouvelle centrifugation pour éliminer toute trace de silice. Le plasmide ainsi obtenu est conservé à -20°C pour utilisation future.

#### VI.6.2. Chez *C. difficile*

Il est également possible d'extraire des plasmides d'une souche de *C. difficile* mais le rendement de cette extraction est très faible et est donc un dernier recours. Il est beaucoup plus profitable de réaliser les extractions plasmidiques à partir des souches d'*E. coli* comme décrit précédemment. Il s'est néanmoins révélé nécessaire d'extraire un plasmide depuis *C. difficile* afin de transformer une *E. coli* TOP10 par la suite.

Pour cela, une culture O/N de 10 mL de *C. difficile* a été centrifugée 1 min/16200 et au culot ont été ajoutés 400 µL de tampon 1 (60 mg/mL de lysozyme, 500 mM de sucrose) avant d'être incubés 30 min à 37°C. Après ajout de 800 µL de tampon 2 (200 mM NaOH, 3% SDS), le mélange est vigoureusement agité puis incubé 5 min à température pièce d'ajouter 600 µL de tampon P3 (le même que chez *E. coli*) et de mélanger par inversion. Le tout est ensuite mis sur glace 5min avant d'être centrifugé 5 min/16200g. Le surnageant est récupéré et additionné à 800µL de tampon QG (Quiagen) et à 200 µL de silice fraîchement resuspendue (cette dernière sédimente rapidement). Après 3 min d'incubation à température ambiante en resuspendant toutes les minutes, le mélange est centrifugé 10 sec/16200g puis laissé à décanter. La silice est alors lavée avec 500 µL de tampon de lavage (voir section précédente) deux fois puis de nouveau centrifugée 10 sec/16200g. Le culot est alors séché 15min à l'air libre avant d'être resuspendu dans 50 µL de Tris-HCl 10 mM pH 8 via incubation 2 min à 55°C. De même que chez *E. coli*, cette étape est répétée pour éliminer les traces de silice. Les plasmides sont alors prêts à être digérés, transformés ou séquencés.



## VI.7. ClosTron

La technologie ClosTron (clostron.com) a été spécialement conçue pour les *Clostridium* par l'équipe de Nigel Minton (Heap et al., 2007, 2009). Le plasmide (pMTL007C-E2 donné comme exemple dans la Fig. 10 et en annexe 1 pour *grdX*) comprend un intron de groupe II (en jaune) dans lequel un gène codant pour la résistance à l'érythromycine (*ermB*) est inséré. Toutefois, cette résistance n'est pas fonctionnelle car au sein du gène est inséré un intron de type I (en noir) provenant du phage td. Cet intron peut contrôler son propre épissage du transcrit d'ARN mais ce mécanisme nécessite une orientation spécifique. Normalement, le gène *ltrA* (en vert) est présent au sein de l'intron de groupe II mais a été déplacé sur le plasmide. Lorsque transcrit, l'ARNm contenant le gène *ermB* présente toujours l'intron td de type I car l'orientation de la transcription ne permet pas l'épissage auto-catalytique de ce dernier. Le gène *ermB* étant toujours interrompu, la bactérie ne possède pas de résistance à l'érythromycine (Em) et y est donc sensible (Em<sup>S</sup>).

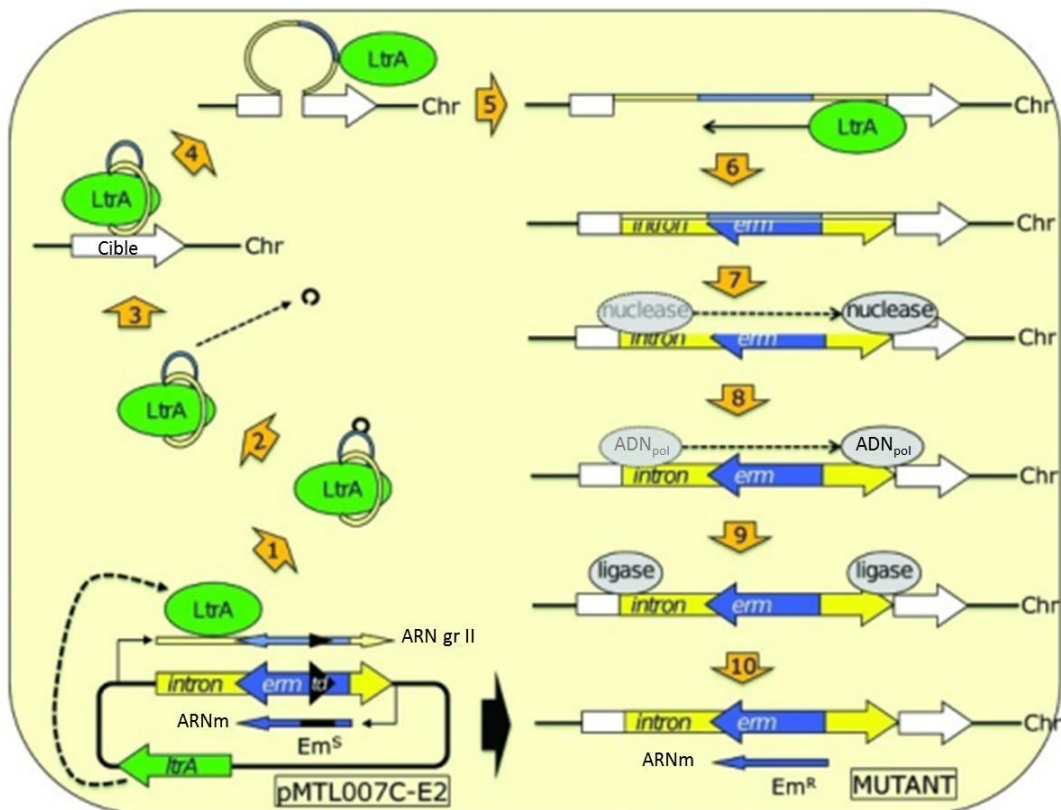


Figure 17. Principe du ClosTron. Adapté de (Kuehne and Minton, 2012)

Toutefois, lorsque le brin opposé d'ADN contenant l'intron de groupe II est transcrit, la protéine LtrA peut venir se lier au transcrit pour former un complexe protéique ribonucléique (CPR) (1) et comme l'intron se trouve à présent dans la bonne orientation, il s'épisse (2). Le CPR reconnaît et lie ensuite des séquences spécifiques dans le gène cible au sein du chromosome (Chr) (3). Cette reconnaissance est rendue possible grâce à la séquence dégénérée, faite pour reconnaître le gène cible, insérée dans l'intron de groupe II lors de la construction du plasmide. La protéine LtrA vient alors couper dans le gène cible et insère l'ARN (4). L'activité de rétrotranscriptase de cette protéine permet la synthèse du brin complémentaire d'ADN (5) formant brièvement un hybride ADN/ARN (6) avant que des nucléases de la bactérie ne viennent dégrader l'insert d'ARN (7) qui sera remplacé par de l'ADN via une ADN polymérase (ADN<sub>pol</sub>) (8). La ligation des brins est assurée par des ligases bactériennes (9) ce qui résulte en un gène *ermB* fonctionnel (10) qui confère la résistance à l'érythromycine à la bactérie (Erm<sup>R</sup>) et permettra de sélectionner les mutants.

Lors du design de l'intron de groupe II (visant à lui faire reconnaître une séquence précise sur le chromosome bactérien) à l'aide de l'outil bio-informatique disponible sur le site internet (compagnie DNA2.0), plusieurs possibilités de sites d'insertion sont données. Il a toujours été privilégié des insertions en début du gène visé afin de limiter la possibilité de production de protéines tronquées qui pourraient avoir un effet dans la bactérie et ainsi biaiser les résultats obtenus.

## VI.8. Transformation et conjugaison

### VI.8.1. Transformation

Une fois le plasmide construit, il doit être introduit dans une souche d'*E. coli*. Pour les TOP10, 1 µL de réaction de ligation ou de Gibson est ajouté à 50 µL de cellules qui sont ensuite mises sur glace pendant 30min. Un choc thermique de 45 sec à 42°C en bain-marie est réalisé puis les tubes sont remis sur glace pendant 2 min avant d'ajouter 250 µL de LB. Les tubes sont alors incubés à 37°C et agités à 220 rpm pendant 1h à la suite de laquelle 100 µL sont déposés sur gélose LB supplémentée de 25 µg/ mL de chloramphénicol (Cm25). L'homogénéité du dépôt est assurée par l'ajout de billes de verre sur la gélose et son agitation. Les boîtes de Pétri sont ensuite incubées O/N à 37°C. Les contrôles usuellement utilisés comprennent le plasmide non-digéré avec les bactéries

(contrôle de transformation), les bactéries seules (contrôle de l'efficacité de la sélection de l'antibiotique pour le plasmide) et le plasmide digéré sans insert (contrôle du bruit de fond). Les colonies positives sont ensuite contrôlées par PCR ou par extraction du plasmide suivie d'une digestion enzymatique visualisée sur gel d'agarose 0,8%. Si les résultats sont concluants, les plasmides (insert plus jonctions sont envoyés à séquencer au CHUS de Laval avant d'être transformés selon le même protocole (avec addition de kanamycine 100 µg/mL dans les géloses pour maintenir le plasmide conjugatif R702)) dans la souche CA434 en vue de la conjugaison avec *C. difficile*.

Pour la souche *B. subtilis* 168, environ 200 ng de plasmide sont ajoutés à 500 µL de cellules compétentes avant d'être incubés 30 min à 37°C au bout desquelles 300 µL de LB sont adjoints. Après une nouvelle étape d'incubation dans les mêmes conditions, 150 µL sont déposés sur gélose LB supplémentée de 5 µg/mL de chloramphénicol puis incubés O/N à 37°C. L'intégration du plasmide est ectopique dans le génome.

#### VI.8.2. Conjugaison

La conjugaison s'effectue entre une souche d'*E. coli* CA434 pourvue d'un plasmide mobilisable et une souche d'intérêt de *C. difficile* 630. Les deux espèces bactériennes sont mises en culture pour la nuit dans : du bouillon LB supplémenté en chloramphénicol (25 µg/mL) et en kanamycine (100 µg/mL) à 37°C/220 rpm en aérobie pour l'*E. coli* tandis que le *C. difficile* est lui incubé dans du bouillon Cerveau-Cœur supplémenté en cystéine à 1 mg/mL (BHIS pour *Brain-heart infusion supplemented*) à 37°C en anaérobie.

Le jour suivant, la bactérie donneuse (*E. coli* CA434) est centrifugée à 2000g pendant 2 min et le surnageant est ôté. Le culot est alors lavé très doucement avec 1 mL de PBS puis centrifugé de nouveau pour enlever le surnageant. Les cellules sont ensuite resuspendues dans 200µL de culture de *C. difficile* et ce, toujours très doucement pour ne pas endommager les pili nécessaires à la conjugaison. Le mélange de culture est alors déposé goutte-à-goutte sur une gélose BHIS-Agar sans antibiotique et incubé 8 à 24h à 37°C en anaérobie sans retourner la boîte de Pétri.

Le lendemain, les transconjugants sont récupérés à l'aide d'un râteau stérile en grattant la gélose avec 2 mL de PBS dégazé. La suspension est alors étalée sur gélose BHIS-Agar supplémentée avec 12 µg/mL de norfloxacine (utilisée pour tuer les *E. coli*)

et 15 µg/mL de thiamphénicol pour le maintien du plasmide chez *C. difficile*. Les cultures des deux bactéries du jour précédemment sont également déposés sur Pétri comme contrôle. Les géloses sont ensuite incubées en anaérobiose à 37°C pendant 24 à 72h jusqu'à l'obtention de colonies.

Lorsque celles-ci sont obtenues, elles sont repiquées sur des géloses de sélection contenant 2,5 µg/mL d'érythromycine puis incubées à 37°C en anaérobiose pendant au moins une nuit. Plusieurs repiquages peuvent s'avérer nécessaires pour diminuer le risque de faux positifs. Les colonies sont analysées par PCR pour vérifier : 1) l'intégration de la cassette ainsi que la présence de l'intron II, 2) l'excision de l'intron I, 3) le sens d'insertion de l'intron II. Pour se faire, une colonie est resuspendue dans 50 µL de Chelex (Sigma-aldrich) 5% stérile qui est un chélatant venant lier le matériel polaire de la bactérie, soit les débris, tout en protégeant l'ADN lors des étapes de chauffage. En effet, la suspension est ensuite incubée 10 min à 98°C puis centrifugée à 16200g pendant 1 min. Le surnageant est alors utilisé pour les PCR.

Les mêmes amorces servent à confirmer la présence de l'intron II ainsi que l'excision de l'intron I selon la taille du fragment obtenu (1286 pb contre 894 pb, l'intron I faisant 392 pb). Le sens d'insertion est quant à lui déterminé en utilisant une des amorces ci-dessous (présentes dans l'intron) avec une amorce dans la direction inverse située dans le gène muté.

Tableau 13. Amorces utilisées pour confirmer la présence de l'intron II et l'excision de l'intron I.

Nom	Séquence 5'-3'	Direction
LCF-442	CGACGCGTGCGACTCATAGAATTA	Sens
LCF-443	GCTATCATTGCCATTTCCTAACGC	Anti-sens

## VI.9. Courbes de croissance

Les courbes de croissance ont été réalisées en milieu riche, *Tryptone Yeast* (TY) ainsi qu'en milieu minimum (MM) (Karlsson et al., 1999). Le milieu minimum est un milieu pauvre, dépourvu entre autres de glycine. Sa composition est la suivante : tryptophane, 100 mg/mL ; Méthionine, 200 mg/mL ; Isoleucine, 300 mg/mL ; Proline, 300 mg/mL ; Valine, 300 mg/mL ; Leucine, 400 mg/mL ; Cystéine, 500 mg/mL ; CoCl<sub>2</sub>·6H<sub>2</sub>O, 1 mg/mL ; FeSO<sub>4</sub>·7H<sub>2</sub>O, 4 mg/mL ; MnCl<sub>2</sub>·4H<sub>2</sub>O,

10 mg/mL ;  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ , 20 mg/mL ;  $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$ , 26 mg/mL ;  $(\text{NH}_4)_2\text{SO}_4$ , 40 mg/mL ;  $\text{KH}_2\text{PO}_4$ , 0,9 g/mL ;  $\text{NaCl}$ , 0,9 g/mL ;  $\text{NaHCO}_3$ , 5 g/mL ;  $\text{Na}_2\text{HPO}_4$ , 5 g/mL ; Glucose, 2 g/mL ; D-Biotine 12  $\mu\text{g/mL}$  ; Calcium D-Pantothénate 1 mg/mL ; Pyridoxine 1 mg/mL. Le milieu est préparé puis filtré à 0,2 $\mu\text{m}$  avant d'être dégazé pendant la nuit avant utilisation comme le TY.

La veille, 5 mL de TY dégazé sont inoculés à partir du stock congelé et sont incubés durant la nuit à 37°C. Si la souche contient un plasmide (généralement le pRPF144), 15 mg/mL de thiamphénicol sont ajoutés pour maintenir la présence du plasmide. La culture est centrifugée au matin et resuspendue dans du milieu minimum puis utilisées pour inoculer en triplicatas 5 mL de TY ou de MM (avec ou sans thiamphénicol) afin d'obtenir une  $\text{DO}_{600}$  égale à 0,1. Les cultures sont ensuite placées en anaérobiose à 37°C entre les mesures de DO survenant toutes les 1 à 2h après une brève agitation.

#### VI.10. Essais $\beta$ -glucuronidase chez *C. difficile*

Pour ces essais rapporteurs, le riborégulateur est placé en amont du gène *gusA* codant pour la p-glucuronidase, inséré via les enzymes de restrictions EcoRI et KpnI (Annexe 2). Cette enzyme catalyse la transformation du 4-nitrophényl-p-D-glucuronide en acide p-D-glucuronique et en p-nitrophénol donnant une coloration jaune au milieu et qui sera suivie par absorbance à 420nm (Fig. 18).

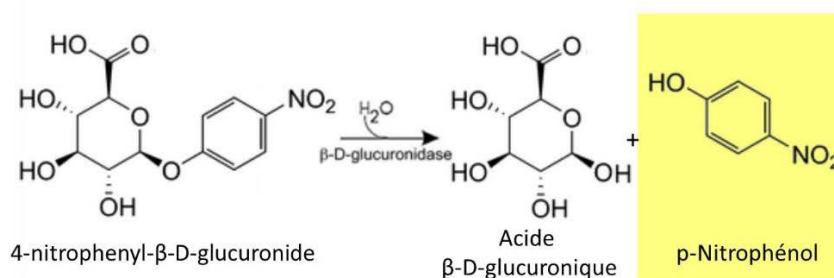


Figure 18. Réaction catalysée par la  $\beta$ -D-glucuronidase.

Ce gène est présent sur le plasmide pRPF144 (voir carte en supplémentaire 1) et est précédé de son promoteur constitutif  $P_{\text{cwp2}}$  qui est utilisé dans les constructions. Le riborégulateur est amplifié par PCR avec des amorces contenant des sites de restrictions

puis digéré et inséré par ligation dans le plasmide lui-même digéré. Le clonage et la transformation s'effectuent dans la souche *E. coli* TOP10 où les bactéries contenant le plasmide sont sélectionnées sur gélose LB supplémentée en chloramphénicol à 15 µg/ mL. Les transformants sont ensuite confirmés par extraction plasmidique (voir VI.4.) puis digestion enzymatique pour visualisation de l'insert et du plasmide. Des produits PCR dont l'amplification comprend le riborégulateur ainsi que des séquences de part et d'autre sur le plasmide sont envoyées à séquencer pour confirmation finale. Le plasmide est alors transféré par conjugaison dans la souche de *C. difficile* 630Δ*erm* où la pression de sélection est maintenue via l'ajout de thiamphénicol 15 µg/mL.

#### VI.10.1. Essais β-glucuronidase sur extraits cellulaires

Le protocole utilisé ici a été développé au laboratoire Fortier pour le riborégulateur guanine RR-guaA. Des colonies isolées de la souche d'intérêt avec le plasmide pRPF144 sont inoculées dans 10 mL de bouillon TY supplémenté avec 12 µg/mL de norfloxacine et 15 µg/mL de thiamphénicol puis incubées O/N à 37°C en anaérobiose. Le lendemain, du milieu frais est inoculé à 2% et incubé dans les mêmes conditions jusqu'à obtenir une DO<sub>600</sub> de 0,7. Les cultures sont alors centrifugées à 3700g/8min puis lavées avec du PBS 1X dégazé avant d'être resuspendues dans 5 mL de milieu minimum (le même qu'utilisé pour les courbes de croissance) supplémenté avec 15 µg/mL de thiamphénicol. Les cultures sont alors incubées à 37°C jusqu'à obtenir une DO de 0,7 à laquelle des concentrations croissantes de glycine sont ajoutées puis incubées 2h à 37°C. À l'issue de cette incubation, la DO est notée et 1 mL de culture est transféré dans un tube Eppendorf avant d'être centrifugé à 3700g/2 min puis lavé au PBS 1X. Le culot est ensuite resuspendu dans 500 µL de tampon Z (100 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7 ; 40 mM NaH<sub>2</sub>PO<sub>4</sub>, H<sub>2</sub>O ; 10 mM KCl ; 1 mM MgSO<sub>4</sub>·7H<sub>2</sub>O et 50 mM 2-mercaptoéthanol) ou conservé à -80°C. Ce tampon contient des cofacteurs et contribue à la stabilité de l'enzyme.

Une fois les culots solubilisés, 100 µg/mL d'endolysine sont ajoutés et la suspension est incubée 1h à 37°C jusqu'à ce qu'elle devienne limpide. Elle est alors mise sur glace pendant 5 min puis centrifugée à 3700g/2 min puis 100 µL de lysats sont prélevés et déposés dans une plaque 96 puits. Après une incubation de 5 min à 37°C, la réaction enzymatique est amorcée par l'ajout de 40 µL de p-nitrophényl-β-D-glucuronide 6 mM. La DO est alors mesurée toutes les 2 min à 405 nm.

Pour rapporter l'activité à la quantité de totales de protéines solubilisées, une courbe standard de Bradford est réalisée. Pour cela, des aliquotes de BSA de 0,05 à 0,5 mg/mL sont préparés dont 10 µL sont ajoutés à 200 µL de colorant de Bradford dilué 1:3 et incubés 30 min. La DO est alors mesurée à 595 nm et la quantité de protéines reportée sur une courbe étalon. Les échantillons de l'essai β-glucuronidase sont testés de la même façon.

L'activité enzymatique est alors calculée de la manière suivante puis rapportée à la quantité de protéines totales déterminée par Bradford:

$$\text{Activité spécifique} = \frac{1000 \times DO_{405nm}}{\text{Volume en mL} \times \Delta t \text{ en min}}$$

#### VI.10.2. Essais β-glucuronidase sur cellules perméabilisées

Un autre protocole utilisant des cellules perméabilisées a également été testé et de ce fait le protocole d'essais β-glucuronidase a subi quelques changements. Les cultures sont de 15 mL mais le début reste similaire bien que les culots soient resuspendus dans 2 mL de Z-Buffer. Aux 500 µL prélevés ensuite sont ajoutés 100 µL de SDS 0,1% et 200 µL de chloroforme. Après un court vortex et décantation, 400 µL sont prélevés dans la phase supérieure et incubés 5 min à 37°C avant l'ajout de 100 µL de substrat à 15 mM. Les tubes sont alors incubés 30 min à 37°C puis la réaction est arrêtée par l'ajout de 100 µL de Na<sub>2</sub>CO<sub>3</sub> 1 M avant une centrifugation de 9600g/2 min. La DO est ensuite lue à 405 nm.

L'activité de l'enzyme est calculée de la manière suivante :

$$\text{Activité} = \frac{1000 \times DO_{405nm}}{DO_{600nm} \times \text{Volume en mL} \times \Delta t \text{ en min}}$$

#### VI.11. Essais β-galactosidase chez *B. subtilis*

Les essais β-galactosidase sont un autre type d'essais rapporteurs qui a été utilisé chez *B. subtilis*. Dans ce cas-ci, le riborégulateur est placé en amont du gène *lacZ* codant pour une β-galactosidase. Cette enzyme catalyse la transformation de l'ONPG en galactose et en o-nitrophénol donnant une coloration jaune à la solution (Fig. 19). Le gène



est présent sur le plasmide pDG1661 dans lequel on transfère le riborégulateur entre les sites de restrictions SacI et BamHI avant de transformer *B. subtilis* 168 avec. Le plasmide sera alors intégré au génome.

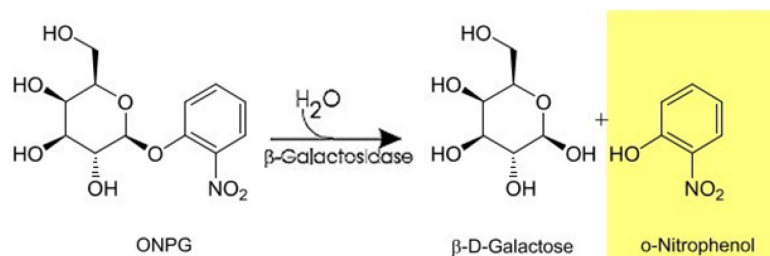


Figure 19. Réaction catalysée par la β-galactosidase.

À partir de cultures de *B. subtilis* inoculées la veille dans du LB et incubées O/N à 37°C, une culture de 150 mL de LB pour chaque souche estensemencée et incubée à 37°C avec agitation jusqu'à atteindre une DO<sub>600</sub> de 0,1. Ces cultures sont centrifugées 10 min/3700g/4°C avant d'être resuspendue dans le même volume d'un milieu minimum. Pour 1L, ce milieu est composé de 200 mL de sels 5X (75,7 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> ; 400 mM K<sub>2</sub>HPO<sub>4</sub> et 220 mM KH<sub>2</sub>PO<sub>4</sub>), 1,7 mL de citrate de sodium 1M, 400 µL de mgSO<sub>4</sub> 1 M ; 4 mL de glutamate 50% ; 5 mL de L-tryptophane 10 mg/mL ; 1 mL de L-lysine 50 mg/mL ; 5 mL de L-méthionine 10 mg/mL ; 8 mL de glucose 50%). Ces 150 mL sont divisés en 6 erlenmeyers de 25 mL, 3 avec glycine (C<sub>f</sub>=10 mM) et 3 sans glycine et la DO est prise. Les cultures sont incubées pendant 6h à 37°C avec agitation et leur DO est notée. Dans chaque erlenmeyer, 8 mL sont prélevés puis centrifugés 10 min/3700g/4°C. Le culot est ensuite lavé avec 800 µL de tampon Tris-HCl 25 mM, pH7,4 et centrifugé 10 min/3700g/4°C. Les culots sont ensuite congelés dans de l'azote liquide et conservés à -20°C.

Les culots sont décongelés sur glace puis resuspendu dans 640 µL/mL de tampon Z avec 40 µL/mL de lysozyme à 10 mg/mL. Ils sont ensuite incubés 5 min à 37°C puis 8 µL de Triton X100 sont ajoutés. Les tubes sont placés 5min sur glace avant d'être préchauffés 5 min à 30°C. À cela, 100 µL d'ONPG à 4 mg/mL - le réactif de la β-galactosidase - sont ajoutés puis les tubes sont placés à l'obscurité jusqu'à ce qu'une coloration jaune apparaisse. La réaction est alors arrêtée par l'ajout de 400 µL de Na<sub>2</sub>CO<sub>3</sub> 1 M et le temps nécessaire à l'apparition de la couleur jaune est noté (Δt). Les tubes sont ensuite centrifugés 5 min/V<sub>max</sub> et les DO à 420 et 550 nm sont prises. Le blanc est constitué de



640 µL de tampon Z, de 200 µL d'ONPG et de 400 µL de Na<sub>2</sub>CO<sub>3</sub> 1 M. L'activité a ensuite été calculée avec la formule ci-dessous et exprimée en Unités Miller :

$$Activité = \frac{1000 \times DO_{420nm}}{\Delta t \text{ en min} \times DO_{600nm}}$$

Avec la densité optique mesurée à 420 nm et 600 nm et Δt (en minutes) le temps de réaction jusqu'à l'obtention de la coloration jaune (t<sub>j</sub>-t<sub>0</sub>).

#### VI.12. Extraction d'ARN de *C. difficile*

La souche 630Δ*erm* de *C. difficile* estensemencée la veille en bouillon TY à partir du stock congelé puis incubée à 37°C en anaérobie. Le lendemain, 100 mL de TY sont inoculés à 3% à partir de cette culture et incubés jusqu'à atteindre une DO<sub>600</sub> de 0,8. Une fois celle-ci atteinte, la culture est répartie en 6x15 mL et centrifugée 15 min/3700g à 4°C. Le culot cellulaire est ensuite resuspendu dans 500 µL de solution de lyse (0,5% SDS, 20 mM acétate de sodium, 1 mM EDTA, ajustée à un pH de 5,5 et filtrée) puis la suspension bactérienne est transférée dans un tube contenant 250 µL de billes de verre et 500 µL de phénol pH 4. Les tubes sont alors soumis à une lyse mécanique (*Bead beater*, Precellys) avec le programme suivant : 6x10sec/2500rpm. Entre chaque 10 sec, les tubes sont mis sur glace pendant 20 sec. Les tubes sont centrifugés 15 min/16200g à 4°C et le surnageant est transféré dans un nouveau tube avec 500 µL de phénol pH 4. Après un vortex bref, ils sont de nouveaux centrifugés 10 min/16200g/4°C. Le surnageant est de nouveau récupéré et ajouté à 500 µL de phénol-chloroforme-isoamylalcool (PCI) avant d'être centrifugés 5 min/16200g/4°C. Cette dernière étape peut être répétée si l'interface n'est pas propre. Le surnageant est récupéré et précipité avec 2,5 volumes d'éthanol 100% et 1/10 de volume d'acétate de sodium 3 M, pH 5,2 avant d'être mis à -80°C O/N. Le jour suivant, les tubes sont centrifugés 30 min/16200g/4°C, le culot lavé à l'éthanol 70 avant d'être de nouveau centrifugé 10 min/16200g/4°C et laissé séché à l'air libre. Il est ensuite resuspendu dans 30 µL d'eau sans RNase et dosé. L'intégrité de l'ARN est également vérifiée sur gel d'agarose 1%.

Après avoir rassemblé les échantillons, un traitement à la DNase (Invitrogen) est réalisé. Pour cela, 1 µL d'enzyme est ajouté pour 10 µg d'ARN puis le tube est incubé

30 min à 37°C après vortex. Le même volume de PCI est adjoint puis le tout est centrifugé 10 min/16200g/4°C. Au surnageant sont ajoutés 10 µL d'acétate de sodium 3 M, pH5,2, 250 µL d'éthanol 100% ainsi qu'1 µL de glycogène puis placés à -80°C pendant 30 à 60 min pour précipitation. S'en suit une centrifugation de 30 min/16200g/4°C, un lavage à l'éthanol 70 puis une nouvelle centrifugation 5 min/16200g/4°C. Un fois le culot séché à l'air libre, il est resolubilisé dans 30 µL d'eau sans RNase. Les ARN sont dosés (NanoDrop 2000, Thermo Scientific) et vérifiés sur gel d'agarose 1% avant d'être congelés à -20°C ou -80°C.

#### VI.13. Extension d'amorce (*Primer extension*)

L'objectif de l'extension d'amorce est de déterminer le site d'initiation de la transcription (le +1) d'un ARN à partir d'une matrice issue d'un produit PCR (Carey et al., 2013). Les analyses bio-informatiques avaient permis de déterminer un +1 théorique et des amorces ont été dessinées afin d'obtenir une matrice d'une taille inférieure à 300 pb. Au sein de celle-ci, une autre amorce (qui servira de sonde) a été dessinée en anti-sens à une centaine de pb du +1 théorique.

La sonde est préparée à partir de 2 µL de l'amorce en question à 10 µM, 2 µL de tampon T4 PNK 10X, 2 µL de l'enzyme T4 PNK, 2 µL de [ $\gamma$   $^{32}$ P]-ATP et de 12 µL d'eau pour un volume final  $V_f$ =20 µL. Après une incubation de 30 min à 37°C, la solution est transférée dans une colonne G-50 et centrifugée 2 min/1000g pour purification.

La réaction d'extension d'amorce nécessite 40 µg d'ARN extrait de *C. difficile* (voir protocole donné en VI.11) auxquels sont ajoutés 5 µL de sonde marquée, 1 µL de dNTPs 10 mM et de l'eau pour un  $V_f$  de 13 µL par réaction. L'ensemble est incubé 5 min à 65°C pour permettre l'hybridation avant d'être placé sur glace pendant 1 min. A cela sont adjoints 4 µL de tampon 5X, 1 µL de DTT 0,1 M, 1 µL d'inhibiteur d'ARNase (*RNAse Out*) et 1 µL d'enzyme SuperScript III pour un  $V_f$  de 20 µL. Après une incubation de 15 min à 52°C pour permettre la transcription inverse, 1 µL de NaOH 4 M est ajouté et le mélange est transféré à 95°C pendant 2 min. Le volume est ensuite ajusté à 100 µL puis l'ARN est précipité à l'éthanol et au NaOAc 3 M comme décrit précédemment. Le culot est solubilisé dans 4 µL d'eau et 4 µL de bleu de formamide puis conservé à -20°C pour utilisation ultérieure.

L'ARN (4  $\mu\text{L}$  soit environ 20  $\mu\text{g}$ ) est ensuite mis sur gel de séquençage 8% pendant 1h30 à 65 W aux côtés des échelles qui sont préparées par PCR à partir de 2  $\mu\text{L}$  de matrice ADN, 1  $\mu\text{L}$  de sonde marquée, 1,5  $\mu\text{L}$  de tampon (thermopolymérase 5X), 1  $\mu\text{L}$  de triton 3%, 1  $\mu\text{L}$  de polymérase d'ADN Vent exo et 7,5  $\mu\text{L}$  d'eau. Ce mélange est ajouté à raison de 3  $\mu\text{L}$  par réaction à 3  $\mu\text{L}$  de mélange pour chaque nucléotide. Les mélanges de nucléotides nécessaires à la réaction de séquençage sont donnés dans le Tab. 15. Les solutions stock de ddNTPs sont à 5 mM tandis que celles de dNTPs sont à 1 mM.

Tableau 14. Amorces utilisées pour les extensions d'amorces.

RR	Nom	Séquence (5'-3')
<b>Symporteur</b>	Matrice sens (LCF-1085)	CTCATCTTTCTAAGTTTATAATTC
	Matrice anti-sens (LCF-1086)	GATAATCCCTGTCCTTTTGC
	Sonde (LCF-1087)	CCGTAAATTTTACAGTCCTAC
<b>grdX</b>	Matrice sens (LCF-1089)	GAAATATGCAATATCAACATTTGG
	Matrice anti-sens (LCF-1090)	CCCAATTTGTCTTCCCTG
	Sonde (LCF-1091)	GTCTCCCATACTCTTCATCC

La réaction est ensuite réalisée à l'aide d'un thermocycleur selon le programme suivant : dénaturation 95°C/1 min ; hybridation 52°C/1 min et élongation 72°C/1 min pendant 25 cycles. La réaction est ensuite arrêtée via l'ajout de 4  $\mu\text{L}$  de bleu de formamide et mis sur gel aux côtés des réactions d'extension d'amorce.

Tableau 15. Mélanges de séquençage correspondant aux différents nucléotides.

A) Mélange T

ddTTP	2,88 $\mu\text{L}$	0,72mM
dGTP	2 $\mu\text{L}$	0,1mM
dATP	0,60 $\mu\text{L}$	0,03mM
dTTP	0,66 $\mu\text{L}$	0,033mM
dCTP	2 $\mu\text{L}$	0,1mM
10X Seq-buffer	2 $\mu\text{L}$	1X
Eau	9,86 $\mu\text{L}$	

B) Mélange C

ddCTP	1,68 $\mu\text{L}$	0,42mM
dGTP	2 $\mu\text{L}$	0,1mM
dATP	0,60	0,03mM
dTTP	2 $\mu\text{L}$	0,1mM
dCTP	0,82 $\mu\text{L}$	0,041mM
10X Seq-buffer	2 $\mu\text{L}$	1X
Eau	10,9 $\mu\text{L}$	

C) Mélange A

ddATP	3,6 $\mu\text{L}$	0,9mM
dGTP	2 $\mu\text{L}$	0,1mM
dATP	0,60 $\mu\text{L}$	0,03mM
dTTP	2 $\mu\text{L}$	0,1mM
dCTP	2 $\mu\text{L}$	0,1mM
10X Seq-buffer	2 $\mu\text{L}$	1X
Eau	7,8 $\mu\text{L}$	

D) Mélange G

ddGTP	1,44 $\mu\text{L}$	0,36mM
dGTP	0,74 $\mu\text{L}$	0,037mM
dATP	0,60 $\mu\text{L}$	0,03mM
dTTP	2 $\mu\text{L}$	0,1mM
dCTP	2 $\mu\text{L}$	0,1mM
10X Seq-buffer	2 $\mu\text{L}$	1X
Eau	11,22 $\mu\text{L}$	

#### VI.14. Cartographie chimique (*In-Line Probing*)

La cartographie chimique utilise la propriété naturelle de l'ARN de se dégrader en solution. En effet, il est instable et intervient différemment selon la structure de l'ARN. L'ARN simple brin est plus fragile et a tendance à adopter plus facilement une conformation en ligne que l'ARN double-brin qui est lui plus stable. Ainsi une attaque nucléophile dite en ligne, d'où le nom anglais d'*in-line probing*, est nécessaire pour assurer un clivage efficace d'une liaison phosphodiester (Fig. 20 A) (Artsimovitch and Henkin, 2009). Ainsi quand l'oxygène présent en 2' adopte une conformation linéaire et interagit avec le phosphore et l'oxygène en 5' du nucléotide suivant, l'oxygène procède à une attaque nucléophile sur le phosphate ce qui résulte en le clivage de la liaison phosphodiester. La liaison du métabolite (M) à un riborégulateur permet de modifier la conformation de l'ARN et donc de stabiliser des régions ce qui va changer le profil de clivage de celui-ci. Cette stabilisation rigidifie la structure et diminue la probabilité que ces régions puissent se retrouver dans une conformation linéaire et donc d'être clivées (Fig. 20 B). Cette technique permet donc d'identifier des changements conformationnels dus à la liaison du ligand que ce soit le site de liaison lui-même ou des repliements de l'ARN.

L'ARN est tout d'abord transcrit *in-vitro* à l'aide d'un promoteur T7 issu d'*E. coli* (5'-TAATACGACTCACTATAG-3') ajouté sur l'amorce sens lors de l'amplification de la matrice. Pour cela, 10 µL de produit PCR de la matrice d'ADN sont ajoutés à 5 µL d'eau, 10 µL de tampon Frank (40 mM Tris-HCl 1M pH7,9 ; 20 mM MgCl<sub>2</sub> ; 2 mM spermidine ; 1 mM de DTT ; 2 mM NaCl), 20 µL de rNTPs 25 mM, 1 µL de DTT 1 mM, 1 µL de YIPP et 2 µL de RNA polymérase T7. Le mélange est ensuite incubé à 37°C pendant 2h puis précipité par l'ajout de 10 µL de NaOAc 3 M et de 250 µL d'éthanol 100%. Après une incubation de 20 min à 4°C, deux centrifugations à 16200g pendant 20 et 10 minutes respectivement entrecoupées d'un lavage du culot à l'éthanol 70, le culot est solubilisé dans 40 µL d'eau et 40 µL de bleu de formamide.

L'ARN est alors migré sur un gel de polyacrylamide 7% pendant 2h à 18 W. Les bandes d'ARN sont ensuite découpées puis électroéluées dans de l'acétate d'ammonium 8 M pendant 1h à 120V avant d'être de nouveau précipités selon le protocole décrit ci-dessus.

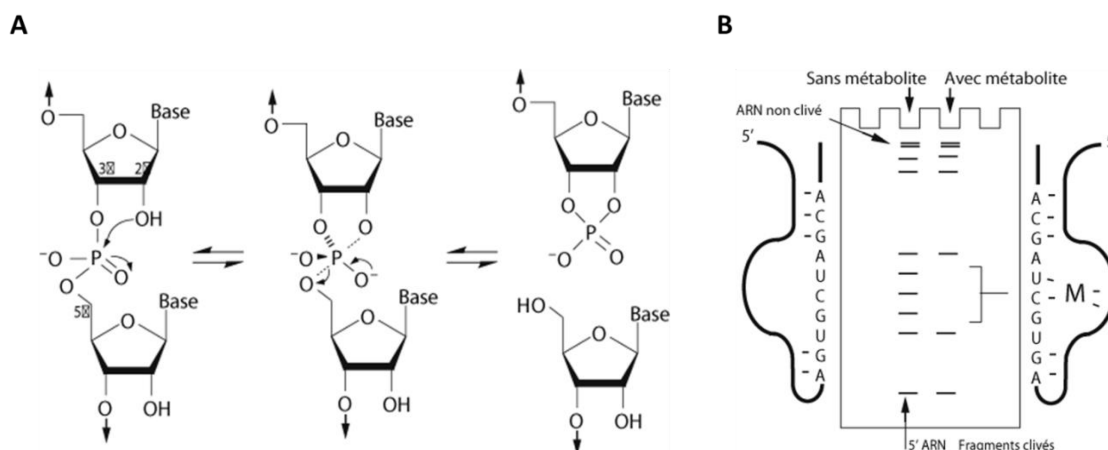


Figure 20. Principe d'une cartographie chimique.

A. Attaque nucléophile en ligne en l'oxygène en 2' et le phosphate adjacent. B. Exemple de résultat d'une cartographie chimique où la liaison du métabolite (M) permet la protection de régions d'ARN du clivage).

Après avoir transcrit l'ARN, il est nécessaire de le déphosphoryler afin de pouvoir le marquer radioactivement. En effet, après une transcription *in vitro*, les ARN sont pourvus d'une extrémité 5' triphosphate qu'il faut enlever. Pour ce faire, 50 pmol d'ARN sont ajoutés à 2  $\mu$ L de tampon 10X, 5  $\mu$ L d'enzyme (*Antarctic phosphatase*) et le mélange est complété avec de l'eau pour un  $V_f$  de 20  $\mu$ L. La solution est ensuite incubée à 37°C pendant 30 min et l'enzyme inactivée après 2 min à 80°C.

La phosphorylation est réalisée en mettant 2  $\mu$ L d'ARN déphosphorylé, 1  $\mu$ L de tampon 10X, 1  $\mu$ L d'enzyme T4 PNK et de l'eau pour un  $V_f$  de 8  $\mu$ L auxquels sont ajoutés 2  $\mu$ L de [ $\gamma$ - $^{32}$ P]-ATP. Le mélange est incubé 1h à 37°C puis la réaction est stoppée par l'ajout de 20  $\mu$ L de bleu de formamide. La totalité de la solution est mise sur gel de polyacrylamide 7% et migrée 1h30 à 18 W.

Les bandes d'ARN marqués sont découpées et mises dans 500  $\mu$ L d'eau à 4°C O/N. Le liquide est ensuite récupéré et l'ARN y est précipité par l'ajout de 900  $\mu$ L d'éthanol 100%, 45  $\mu$ L de NaOAc 3 M et 1  $\mu$ L de glycogène. L'ensemble est placé à -20°C pendant 1h puis centrifugé 30 min/16200g/4°C. Le culot est lavé à l'éthanol 70% puis centrifugé 10 min/16200g/4°C avant d'être solubilisé dans 20  $\mu$ L d'eau.

La réaction de cartographie chimique est réalisée par l'ajout à 1  $\mu$ L d'ARN marqué de 2  $\mu$ L de Tris-KCl 5X, de 2  $\mu$ L de  $MgCl_2$  100 mM, 1  $\mu$ L de glycine à diverses concentrations (de 10 nM à 1 mM) et d'eau pour atteindre un  $V_f$  de 10  $\mu$ L. Le  $MgCl_2$  est ajouté en dernier après une incubation de 3 min à 80°C suivie de 2 min sur glace. La

solution est ensuite incubée 2 jours à température pièce à l'issue desquels 10 µL de bleu de formamide sont ajoutés afin de stopper la réaction.

Les ARN subissent par la suite une électrophorèse sur un gel de séquençage de polyacrylamide 8% pendant 2h30 avec un contrôle d'ARN servant de blanc (-OH), un marqueur de taille moléculaire ainsi qu'un marqueur T1 permettant d'identifier tous les résidus guanine dans la séquence.

#### VI.15. Transcription *in vitro* à cycle unique (*Single Round in vitro transcription*)

La transcription *in vitro* à cycle unique permet d'observer les différences de transcription entre des ARN en fonction des conditions. En effet, étant donné que seule une transcription par copie d'ADN n'est effectuée, il est possible d'observer des différences dans l'intensité relative des transcrits. Par exemple si un ARN est transcrit au complet en absence de ligand, l'ajout de concentrations progressives de celui-ci résultera en la diminution de la quantité d'ARN transcrits complètement au profit d'un transcrit plus court résultant d'une terminaison précoce. C'est cette évolution entre une forme longue correspondant à un transcrit entier et une forme courte issue d'une transcription écourtée qui est recherchée dans cette technique (Artsimovitch and Henkin, 2009).

Cette méthode consiste en l'ajout de 3 nucléotides dont un marqué dans le mélange de transcription (Fig. 21). La polymérase va transcrire la séquence jusqu'à ce qu'elle rencontre le quatrième nucléotide (une douzaine de nucléotides après le promoteur idéalement). N'étant pas en mesure de l'incorporer, la polymérase va s'arrêter ce qui va permettre leur synchronisation. Un nouveau mélange transcriptionnel (appelé pleine longueur) contenant les quatre nucléotides est alors ajouté et la transcription reprend. Une fois celle-ci terminée, la polymérase n'est pas en mesure de recommencer une transcription car elle a été neutralisée par l'héparine.

Pour cela, la matrice d'ADN est amplifiée par PCR avec l'ajout en 5' du promoteur LacUV5 provenant de l'opéron *lac* d'*E. coli* (5'-TAACCAGGCTTTACACTTTATGCT TCCGGCTCGTATAATGTGTGGAAGGATCC-3'). Une fois la matrice obtenue, 1 pmol d'ADN est ajouté à 4 µL de tampon 5X (100 mM Tris-HCl pH 8 ; 100 mM KCl ; 100 mM MgCl<sub>2</sub> ; 500 µM d'EDTA et 70 mM de β-mercaptoéthanol), à 11,7 µL d'eau

ainsi qu'à 1  $\mu\text{L}$  de sigma 70 4  $\mu\text{M}$  et 1  $\mu\text{L}$  de RNAP 2  $\mu\text{M}$ . La solution est ensuite incubée 5 min à 37°C avant d'ajouter le mélange radioactif (1  $\mu\text{L}$  de 2 rNTPs au choix à 25 mM et 0,3  $\mu\text{L}$  de nucléotide marqué). Après 10 min à 37°C, la solution est divisée en deux et chaque moitié est ajoutée à un mélange dit « pleine longueur » avec ou sans ligand et incubé 15 min à 37°C. Le mélange pleine longueur est composé de 2  $\mu\text{L}$  de tampon 5X, 1,5  $\mu\text{L}$  d'héparine 6 mg/mL, 1  $\mu\text{L}$  de rNTPs et 1  $\mu\text{L}$  de ligand ou d'eau puis complété à 10  $\mu\text{L}$  avec de l'eau. À l'issue des 15 min d'incubation, la réaction est arrêtée par l'ajout de 20  $\mu\text{L}$  de bleu de formamide puis l'ensemble est chauffé à 95°C pendant 2 min avant d'être mis sur gel de polyacrylamide 5% et migré pendant 2h30 à 65 W.

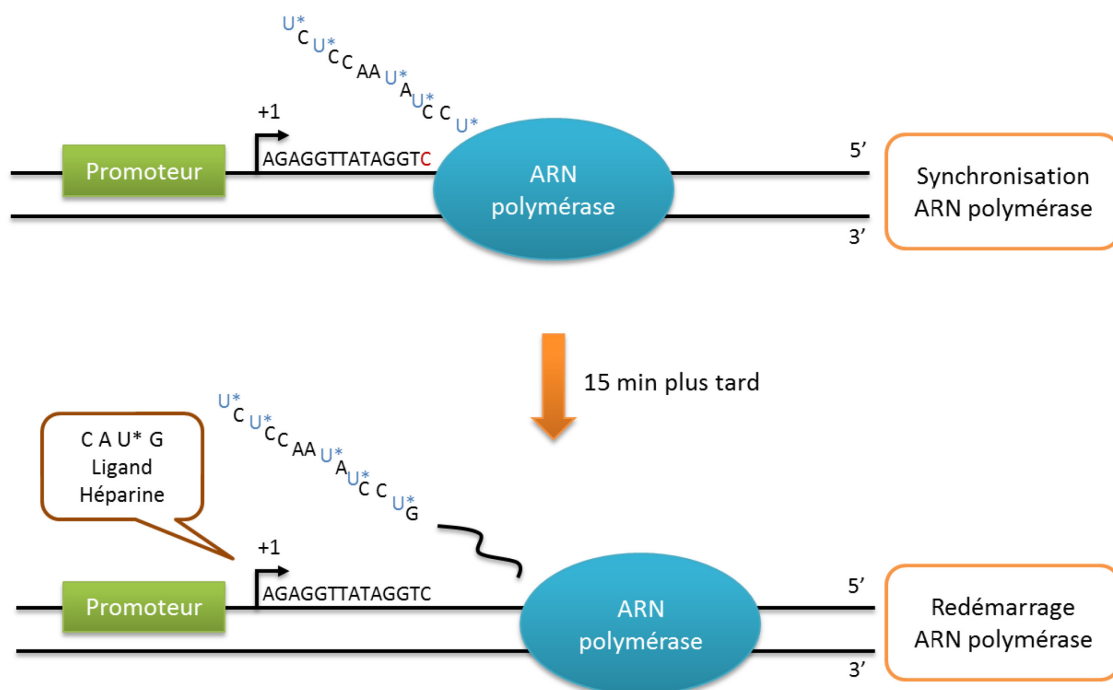


Figure 21. Principe de la transcription in vitro à cycle unique.

Le gel est ensuite séché et exposé O/N avant d'être scanné via un phosphorimager (Typhoon FLA 9500, GE Technologies) et analysé à l'aide du logiciel Quantity One de Bio-Rad Technologies. Les intensités des différentes bandes associées aux différentes tailles de transcrit sont comparées et le T50, point correspondant à 50% de chaque taille de transcrit, est calculé.

## VII. Résultats : Caractérisation du riborégulateur contrôlant le symporteur glycine/sodium chez *C. difficile*

### VII.1. Contexte

Il est relativement commun pour des riborégulateurs glycine de contrôler l'expression de symporteurs sodium/glycine ou sodium/alanine. Chez *C. difficile*, il existe un tel symporteur identifié par bio-informatique, il s'agit du gène CD\_22760 (1368pb). Le riborégulateur avait été lui aussi identifié par analyses bio-informatiques à partir de données d'analyses transcriptionnelles de type RNAseq comme étant un riborégulateur typique en tandem. Il y a pourtant une caractéristique qui le distingue des autres. Il est prédit pour être un riborégulateur dont la liaison du ligand résulte en l'inhibition de l'expression du gène or les riborégulateurs glycine sont connus pour être des activateurs. Cette prédiction est supportée par le fait que la souche (630Lierm) a été cultivée en milieu riche puis les données de séquençage ont été analysées via Rfam. Les riborégulateurs identifiés étaient le plus souvent exprimés et ce en moyenne entre 10 et 100 fois davantage que les gènes qu'ils régulent. Or bien que le riborégulateur qui nous intéresse ici (appelé RR-symp) soit transcrit (Fig. 22), on n'observe aucune expression du gène en aval ce qui suggère que ce riborégulateur en inhibe la transcription.

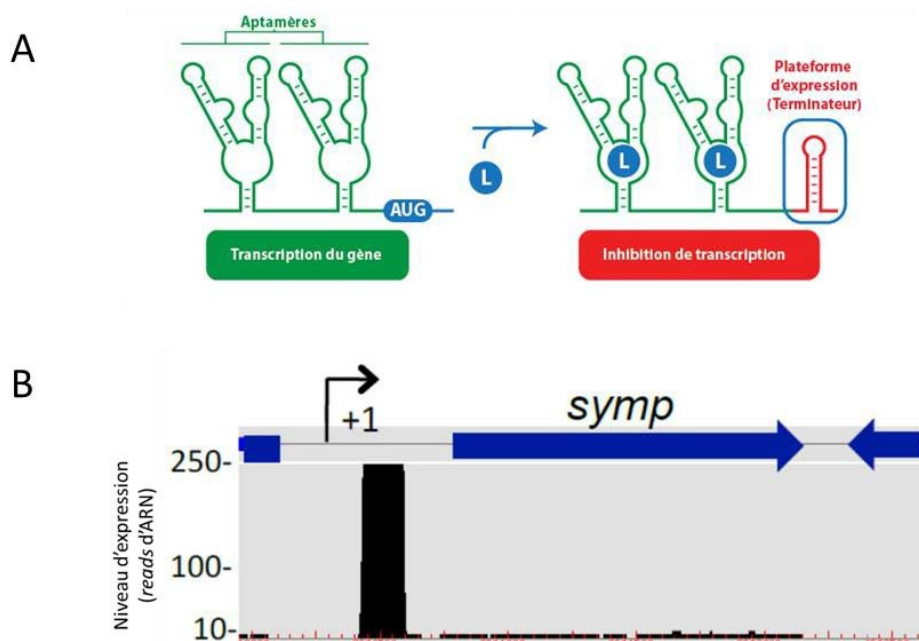


Figure 22. Structure et fonctionnement prédits du RR-symp (A) et données de RNAseq (B).



Il a déjà été évoqué que la glycine pouvait avoir des effets toxiques sur la cellule lorsque les concentrations en cet acide aminé étaient trop importantes (Tezuka and Ohnishi, 2014). Le riborégulateur permettrait donc d'inhiber la synthèse du symporteur pour limiter l'import de la glycine au sein de la cellule. La glycine est importée en temps normal car elle peut être utilisée comme source énergétique. Il existe néanmoins des hypothèses quant au fait que ce symporteur pourrait transporter de l'alanine au lieu de la glycine. Le contrôle de l'expression d'un tel symporteur par la glycine serait expliqué par une voie de biosynthèse de cet acide aminé à partir de l'alanine (Fig. 23). Il s'agirait donc également d'un rétro-contrôle négatif. Néanmoins, de plus en plus d'analyses disponibles sur les sites de voies métaboliques comme KEGG pointent vers un symporteur sodium/glycine plutôt qu'un sodium/alanine.

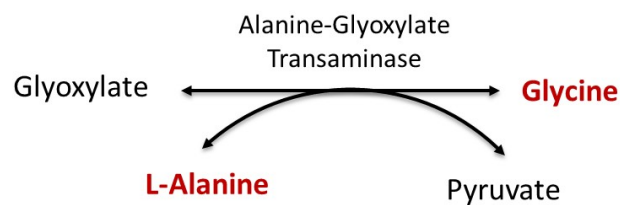


Figure 23. Réaction de la voie de biosynthèse III de la glycine à partir d'alanine.

## VII.2. Résultats

### VII.2.1. Le RR-symp, un riborégulateur avec une structure secondaire typique

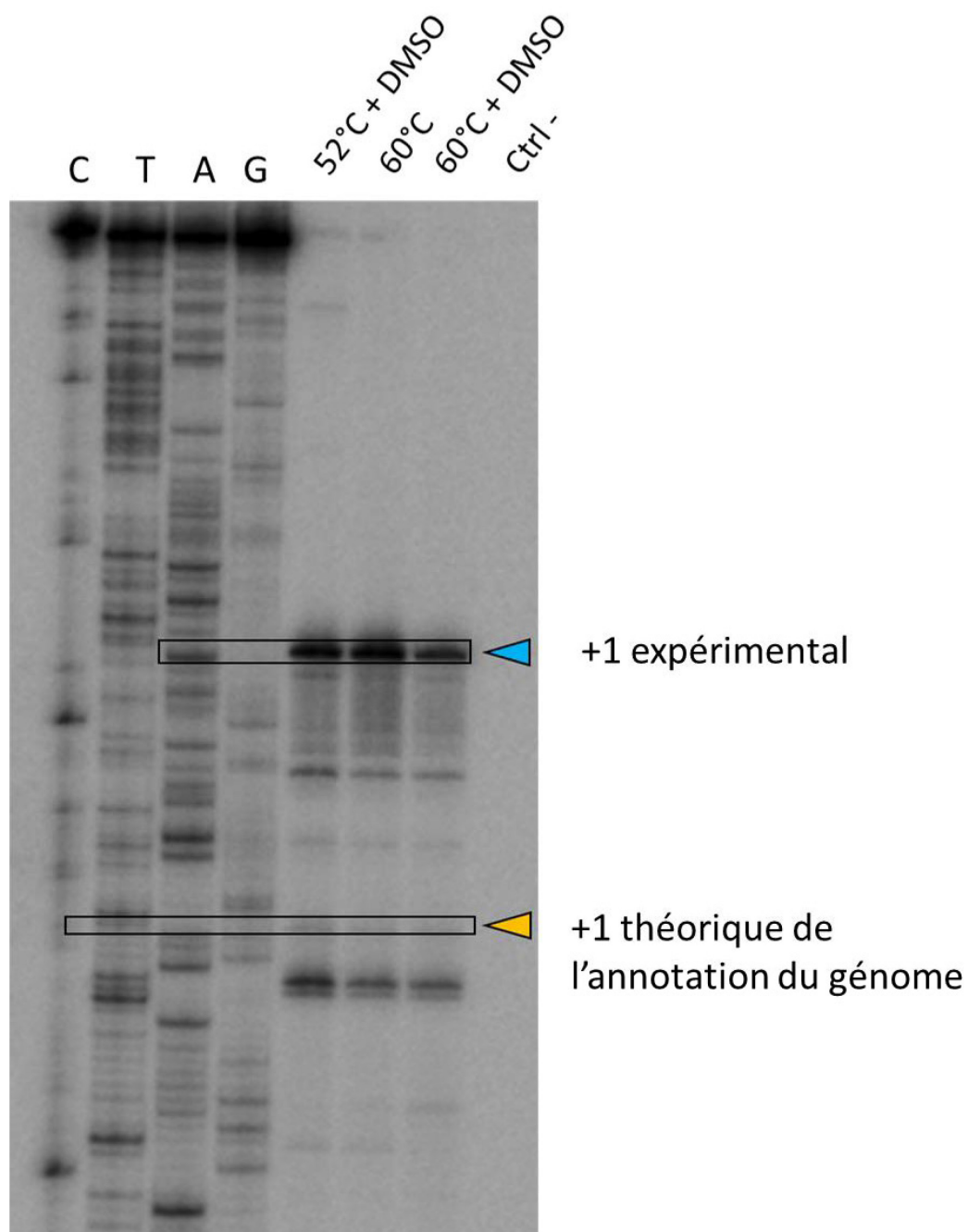
Tout d'abord, il a fallu déterminer si le riborégulateur identifié par bio-informatique était réel et *à fortiori* fonctionnel. Pour cela, il a fallu s'atteler à l'identification de ses structures primaire et secondaire.

Le site d'initiation de la transcription (+1) n'avait auparavant été déterminé que par bio-informatique et une identification plus précise s'avérait nécessaire à la fois pour les essais *in vitro* à suivre mais aussi pour mieux comprendre la structure du riborégulateur en tant que tel. Pour cela, la technique d'extension d'amorces a été utilisée. La matrice d'ADN utilisée couvrait la majeure partie des aptamères du riborégulateur (surligné en jaune sur la Fig. 24) avec une amorce sens située 141pb en amont du site théorique de l'initiation de la transcription (représentée en vert) et une amorce anti-sens à la fin du riborégulateur (séquence soulignée).

Cette expérience a permis l'identification d'une longue séquence en 5' (38pb) du premier aptamère du riborégulateur. Cette séquence ne forme pas de structure secondaire identifiée lorsque la séquence est entrée dans le serveur mFold.

N'étant encore qu'un riborégulateur glycine putatif, il a été important d'en déterminer la structure secondaire et de la comparer avec une structure type. Pour cela, celle du riborégulateur de *V. cholerae* a été utilisée (Fig. 25) et a permis de constater que les structures primaire et secondaire du RR-symp concordent bien avec l'hypothèse d'un riborégulateur glycine en tout cas pour ce qui concerne les aptamères.

Sur la Fig. 24, la matrice a été transcrite et des échelles permettent d'avoir la séquence en nucléotides. Ainsi à l'aide d'une sonde, il est possible de déterminer le nucléotide correspondant au site d'initiation de la transcription.

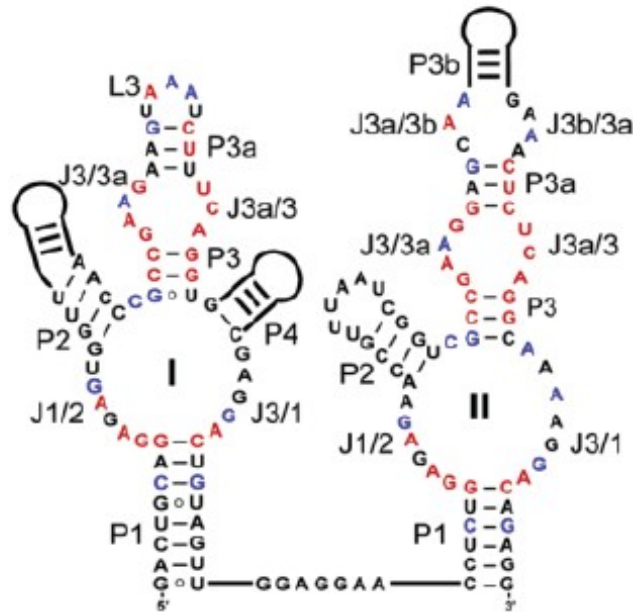


gtatattactcactcttctaaagttataattcgtaaataggcttacatttttatatttttcttgtttttaatttttagttgacaaaac  
 ataaaaacttatattaatattttaataacaataagctatgacaaatcaatatccggtlaagattttatagaagagttgctatata  
 gcacaccgaaggagtaacactctcaggtacttttaaaagtaggactgtaaaatttacggaactctggagagaccat  
 tatgggcgccgaaggggcaaggctttatgctcaatctctcaggcaaaaggacagggattatctacaaaaacagtatt

Figure 24. Extension d'amorce pour le RR-symp et séquence des aptamères (jaune).

La séquence en vert représente l'amorce sens de la matrice tandis que la séquence soulignée indique l'amorce anti-sens de la matrice. La sonde est figurée en gras.

A. *V. cholerae*



B. *C. difficile* symporteur

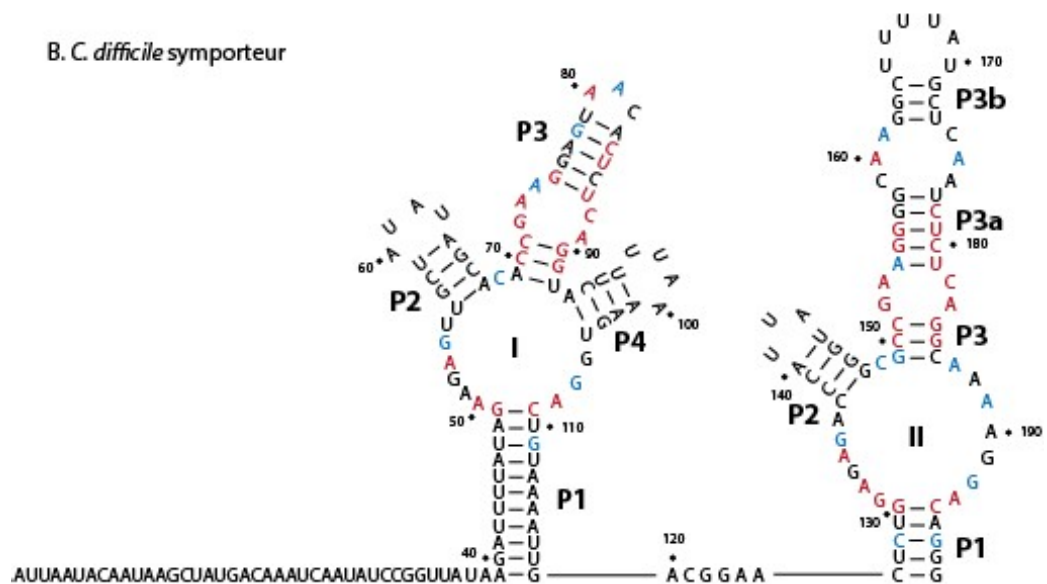


Figure 25. Comparaison entre les structures primaire et secondaire du RR de *V. cholerae* et le RR-symp de *C. difficile*. Les nucléotides en rouge représentent une conservation >95% tandis que ceux en bleu >70%.

## VII.2.2. Une originalité fonctionnelle

Puisqu'il a été montré que ce riborégulateur ressemblait très fortement à un riborégulateur glycine, il a fallu s'assurer que cette structure était bien en mesure de lier la glycine et était fonctionnelle. Pour cela des essais rapporteurs ont été entrepris, d'abord directement chez *C. difficile* puis chez *B. subtilis*.

Le protocole utilisé précédemment dans le laboratoire afin d'obtenir une activité relative à l'expression du gène contrôlé par le riborégulateur s'effectuait directement dans *C. difficile* et utilisait l'enzyme  $\beta$ -glucuronidase. Néanmoins, ce protocole impliquant une lyse des cellules via endolysine n'a pas permis d'observer un effet ligand. L'idée que la méthode de lyse ait pu être responsable de cette absence d'effet a été émise puis testée.

Pour cela, une autre méthode pour accéder à l'activité enzymatique a été adaptée : la perméabilisation des cellules. Si la phase de culture n'a pas été modifiée en dehors des volumes considérés, il a fallu effectuer des tests afin d'optimiser la perméabilisation (Fig. 26). Ces premiers tests furent effectués simplement en termes d'augmentation de la densité optique qui est fonction de la formation du coloré jaune et donc de l'augmentation de la synthèse de l'enzyme et de son activité. Le contrôle positif (Ctrl +) comprend la souche de *C. difficile* avec le plasmide pRPF144 présentant le gène codant pour l'enzyme (*gusA*) mais sans riborégulateur pour en contrôler l'expression. Celle-ci est donc permanente. Le contrôle négatif (Ctrl -) quant à lui présente le plasmide pRPF144E dépourvu du gène *gusA*.

Tout d'abord, le temps d'exposition des cellules au substrat, le p-nitrophenyl- $\beta$ -D-glucuronide à 6 mM comme utilisé dans le laboratoire usuellement a été déterminé. Il s'avère que plus le temps d'exposition est long, plus la réaction est importante (Fig. 26A). Le temps retenu, encadré en vert, fut de 30 min pour les tests ultérieurs. Le volume réactionnel dans lequel ont été resuspendus les culots bactériens avait moins d'influence sur l'expérience mais afin de simplifier les expérimentations et d'économiser du substrat, 0,5 mL a été choisi comme volume final (Fig. 26B). Les expérimentations jusqu'à présent et la littérature utilisaient le substrat de l'enzyme à une concentration de 6 mM mais rien ne garantissait que dans notre cas cette concentration était optimale. Il s'avère qu'en effet, augmenter la concentration en substrat permettait d'obtenir davantage d'activité (Fig. 26C). Un plateau semblant s'établir au-delà de 15 mM, c'est cette concentration qui fut retenue par la suite. La méthode semblait donc fonctionner, restait à déterminer si la souche d'intérêt possédant le plasmide pRPF144 où le gène *gusA* était précédé du riborégulateur RR-symp allait réagir correctement à ce traitement. En effet, la souche d'intérêt réagit de la même manière que la Ctrl + ce qui était attendu étant donné que le RR-symp est un riborégulateur réprimant l'expression du gène (Fig. 26 D). Une dernière variable avant de pouvoir effectuer l'expérience dans des conditions réelles était de s'assurer s'il était possible de congeler les culots avant de procéder à la perméabilisation.

Les phases de cultures pouvant s'avérer longues, il était intéressant de savoir si la congélation à  $-80^{\circ}\text{C}$  O/N avait un quelconque impact sur l'activité et plus précisément un impact négatif. Pour ce faire, la souche Ctrl + a étéensemencée en milieu riche d'une part (TY) et en milieu pauvre d'autre part (MM). Le lendemain, les culots ont été décongelés sur glace et le protocole de perméabilisation a été entrepris selon les conditions préalablement déterminées. Il s'avère qu'un peu d'activité est perdu entre le culot frais et le culot congelé mais pour des raisons pratiques, la congélation a été maintenue dans le protocole. Cette expérience a également été utilisée pour déterminer si le temps d'exposition dans le MM avait une influence. Il s'avère que non, l'activité se maintenant dans le temps, il a alors été décidé de conserver les 2h prescrites dans le protocole initial (Fig. 26E).

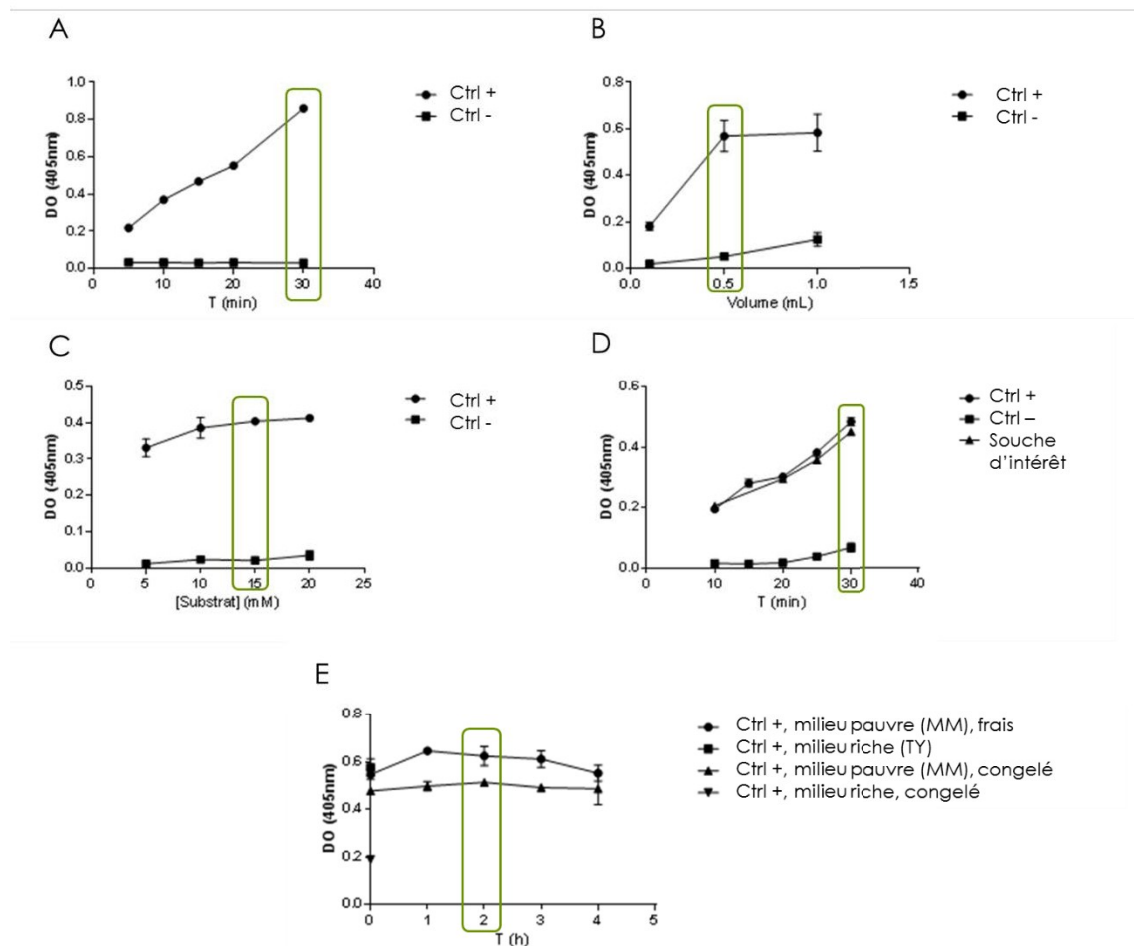


Figure 26. Optimisation de la perméabilisation des cellules pour les essais rapporteurs

A. Détermination de la durée optimale d'exposition au substrat. B. Détermination du volume de cellules optimal. C. Détermination de la concentration optimale en substrat. D. Vérification de la souche présentant le riborégulateur sans ajout de ligand. E. Comparaison de l'activité en fonction de l'état du culot (frais (point unique à T0) ou congelé) et du milieu de culture utilisé (MM ou TY). Les barres d'erreur représentent l'écart-type, N=3 (SD).

L'optimisation étant effectuée, les expérimentations en conditions réelles ont pu commencer. Des concentrations de glycine de 0, 100, 500 et 750 mg/L ont été testées. Mais là encore, comme pour les essais avec l'endolysine, aucun effet ligand n'a pu être observé (Fig. 27).

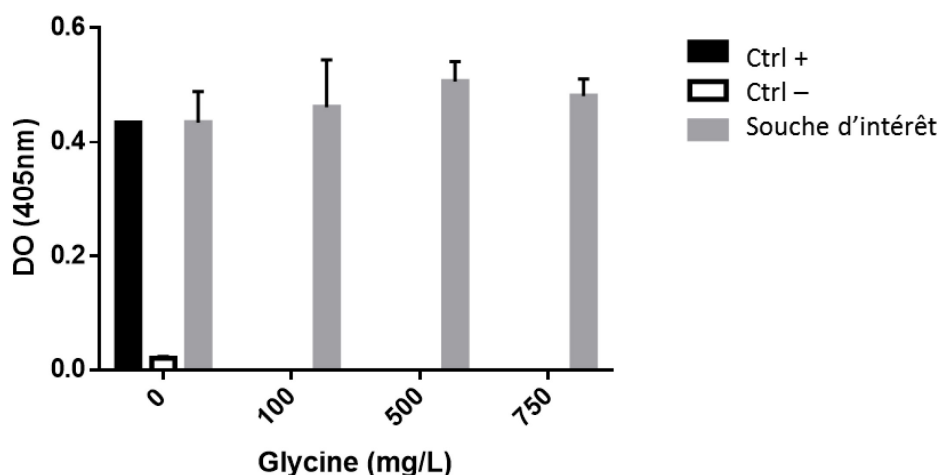


Figure 27. Activité glucuronidase du RR-symp chez *C. difficile* via perméabilisation. N=3, les barres d'erreur représentent l'écart-type SD.

En raison de la difficulté de mettre en place ces essais chez *C. difficile*, il a été décidé que les essais rapporteurs allaient plutôt être conduits chez *B. subtilis*, un autre Gram+, afin de simplifier l'expérience. Ce transfert chez un autre organisme est loin d'être dénué d'intérêt et est retrouvé dans la littérature (Welz and Breaker, 2007). Il est bien sûr plus intéressant de travailler directement dans la bactérie d'origine mais parfois ceci n'est pas possible. Des mutants de *B. subtilis* 168 ont donc été réalisés via le plasmide pDG1661 s'insérant directement dans le génome bactérien. Le riborégulateur a été inséré en amont du gène *lacZ* codant pour la p-galactosidase. Le contrôle négatif est la souche en absence de ligand.

Chez *B. subtilis* 168, on observe effectivement un effet ligand avec une répression de l'activité de l'enzyme indiquant que celle-ci est moins exprimée. L'ajout de glycine entraîne une diminution d'un facteur 1,78 de l'expression en milieu minimum (Fig. 28).

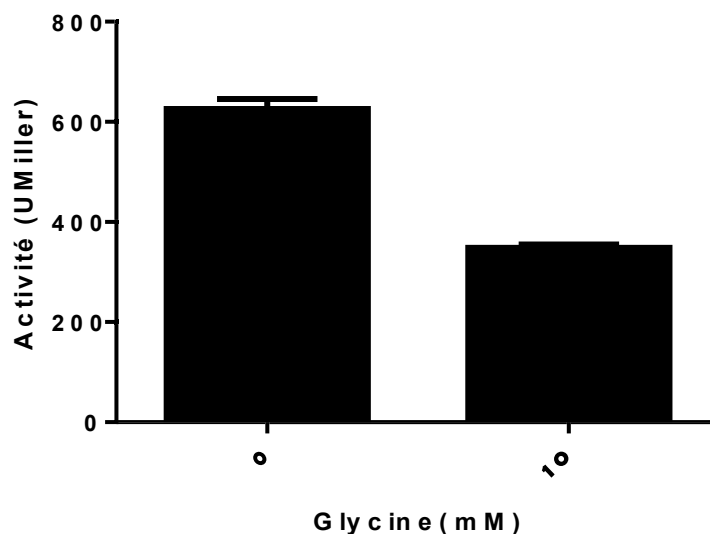


Figure 28. Activité galactosidase du RR-symp chez *B. subtilis*  
N=3, les barres d'erreur représentent l'écart-type SD.

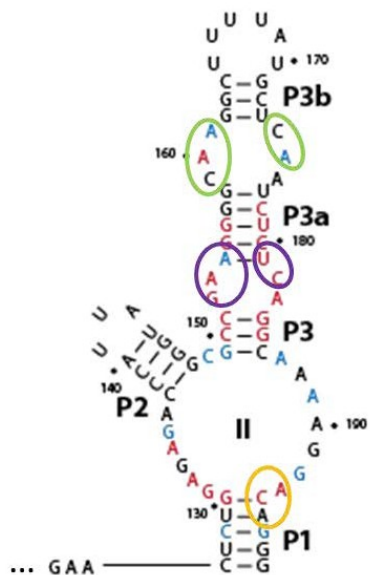
### VII.2.3. Une affinité en accord avec la littérature

Des essais de cartographie chimique (*in line probing*) ont été tentés sur ce riborégulateur que ce soit sur les aptamères au complet ou pris individuellement. Il s'avère que seul le second aptamère a permis l'observation de changement de protection en absence ou en présence de ligand à des concentrations croissantes. Il a ainsi été possible de déterminer plusieurs zones où l'ajout progressif de glycine diminuait la protection de nucléotides. Ces zones sont identifiées par des flèches sur la figure 29. Elles sont retrouvées principalement au niveau des boucles de l'aptamère et ce de part et d'autre de celle-ci. Trois zones sont ainsi isolées : la première (en orange) intervient à l'extrémité de la P1, la seconde (en violet) au niveau du site de liaison du ligand et la troisième (en vert) au niveau de la boucle entre les tiges P3a et P3b intervenant dans le repliement (interaction  $\alpha$ ).

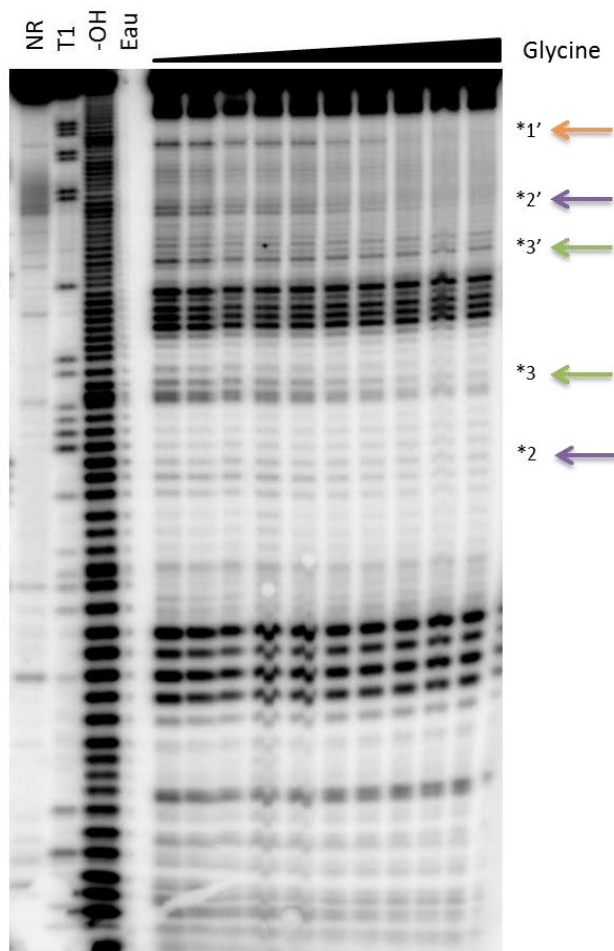
L'étude de ces zones a également permis de déterminer la constante de dissociation ( $K_d$ ), soit l'affinité de l'aptamère pour son ligand. Deux zones les plus fortement impactées par l'ajout de ligand ont pour cela été analysées : la zone 1' et la zone 3'. Dans le cas de la zone 1', le  $K_d$  est de  $4,34 \mu\text{M}$  (Fig. 29) tandis que celle de la zone 3' est de  $3,48 \mu\text{M}$  (graphe non présenté).



## Aptamère 2



G196 ►  
G191 ►  
G184 ►  
G171 ►  
G163 ►  
G158 ►  
G155 ►  
G151 ►  
G149 ►



## A p t a m è r e 2 - Z o n e 1'

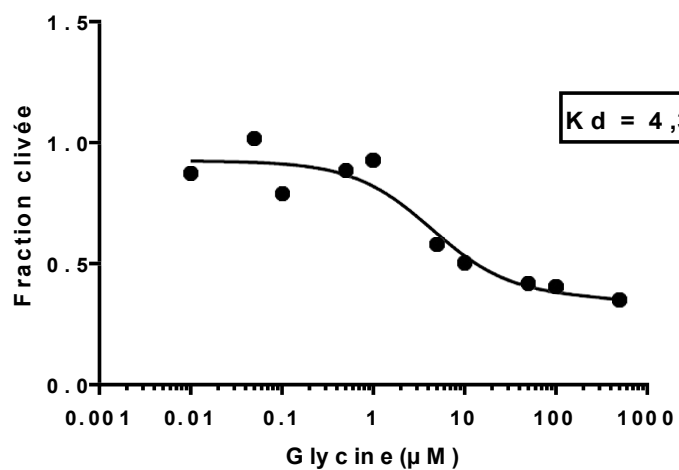


Figure 29. Cartographie chimique de l'aptamère 2 du RR-symp et identification en couleur des zones où le clivage survient associé à une concentration croissante en ligand (haut), T50 sur la zone 1' (bas).

#### VII.2.4. Un mutant à la croissance retardée

Une des premières questions posée fut : l'absence du gène codant le symporteur est-elle délétère pour la bactérie ? Cette interrogation est cruciale pour déterminer si ce riborégulateur pourrait être utilisé comme cible thérapeutique. Pour répondre à cette question, l'inactivation du gène CD\_22760 codant pour le symporteur sodium/glycine (appelé *symp*) a donc été effectuée par la méthode du ClosTron. Brièvement, il s'agit d'insérer un intron au sein du gène visé afin de l'inactiver, insertion couplée avec l'introduction d'un gène codant pour la résistance à l'érythromycine. Une fois le mutant obtenu et confirmé, des courbes de croissance ont pu être effectuées.

En milieu riche, il n'y a pas de différence significative entre les croissances de la souche sauvage d'une part et la souche mutante d'autre part (Fig. 30A). Le mutant croît même légèrement mieux que la souche sauvage.

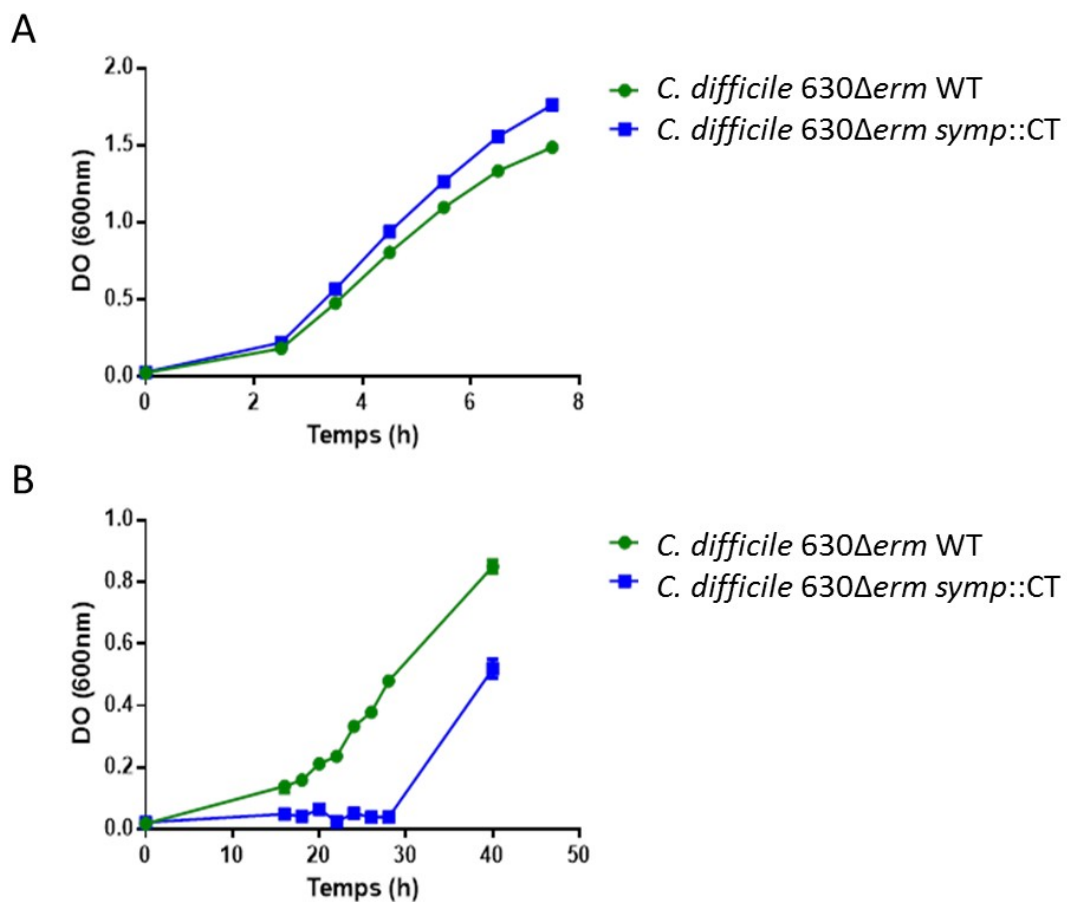


Figure 30. Courbes de croissance de la souche WT et le mutant inactivé du *symp* (*symp*::CT) en milieu riche (A) et milieu pauvre (B). N=3.

En milieu pauvre, on observe un retard de croissance d'une dizaine d'heures pour le mutant *symp::CT* avant que celui n'atteigne une phase exponentielle de croissance (Fig. 30B). Toutefois il a été observé que ce retard est par la suite comblé et les phases stationnaires se rejoignent.

## VIII. Résultats : Caractérisation du riborégulateur contrôlant la glycine réductase *grdX* chez *C. difficile*

### VIII.1. Contexte

Les gènes impliqués dans la dégradation de la glycine sont également largement régulés par les riborégulateurs glycine. Chez *C. difficile*, la glycine réductase est encodée par un ensemble de gènes dont l'enchainement évoque un opéron. La même étude de RNAseq (Soutourina et al., 2013) a permis l'identification d'un autre riborégulateur glycine qui lui se trouve en amont d'un gène codant pour une glycine réductase putative *grdX* (CD\_23570 de 372pb). L'expression du riborégulateur est très importante, entre 0 et 100 fois supérieure à celle du gène *grdX* (Fig. 31). Le gène étant transcrit en milieu riche, il est donc supposé que le riborégulateur *grdX*, appelé RR-*grdX*, permet l'activation de l'expression de son gène en présence de glycine.

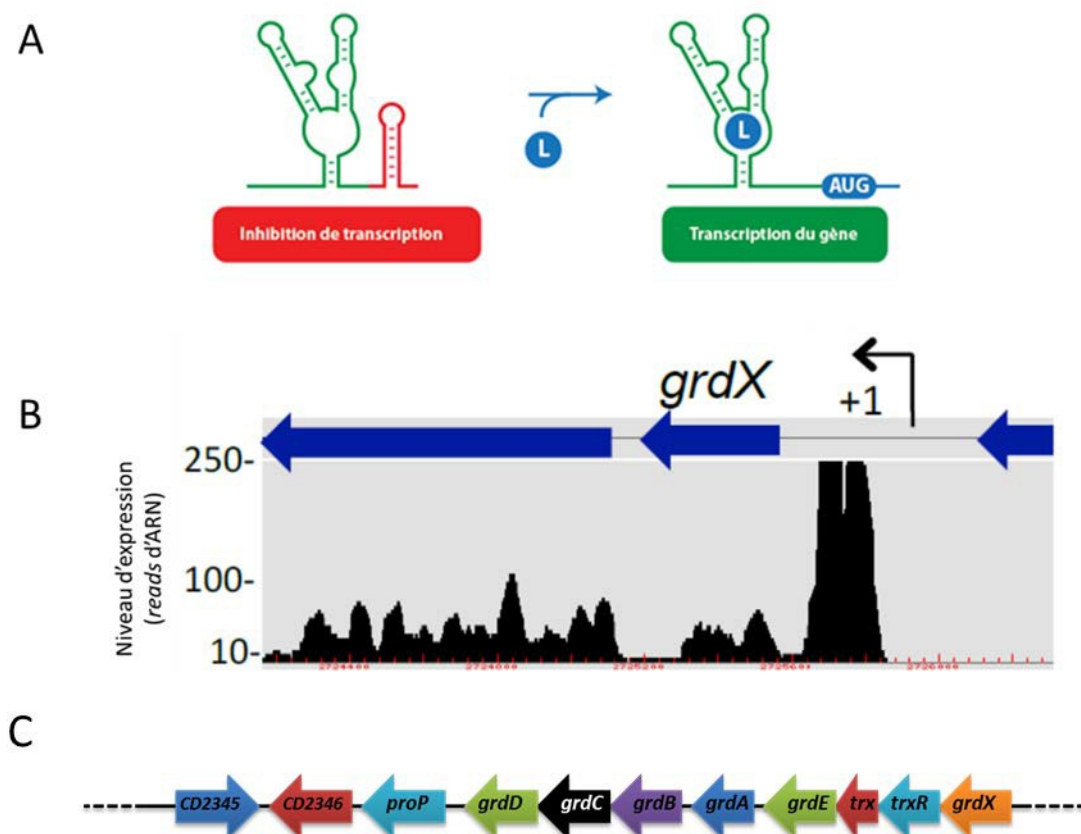


Figure 31. Fonctionnement prédict du riborégulateur *grdX* (A) et données de RNAseq (B), l'opéron potentiel de *grdX* est présenté en (C) où les autres gènes sont également impliqués dans la réductase en codant pour des sous-unités notamment.

En présence d'une concentration importante de glycine issues de l'alimentation de l'hôte, il y aurait donc induction de l'expression de la glycine réductase. Cette enzyme est importante au sein du métabolisme car elle intervient dans la réaction de Stickland permettant la production d'acétate mais surtout d'ATP (Bouillaut et al., 2013; Nisman, 1954) (Fig. 32). Cette réaction d'oxydation/réduction d'acides aminés permet la synthèse d'acides organiques. Un acide aminé sert de donneur d'électron, et est oxydé en un acide carboxylique volatile : par exemple l'alanine en acétate. Un autre acide aminé, la glycine notamment, est réduit pour former également un acide carboxylique, l'acétate ici aussi. Ce mécanisme permet d'éviter l'utilisation d'ions hydrogène comme accepteur d'électron et ainsi de produire de l'hydrogène gazeux. Cette réaction a été observée dans plusieurs *Clostridium* protéolytiques tels que *C. sticklandii*, *C. perfringens*, *C. sporogenes*, *C. botulinum* et *C. difficile* qui nous intéressent ici (Bouillaut et al., 2013; Fonknechten et al., 2010).

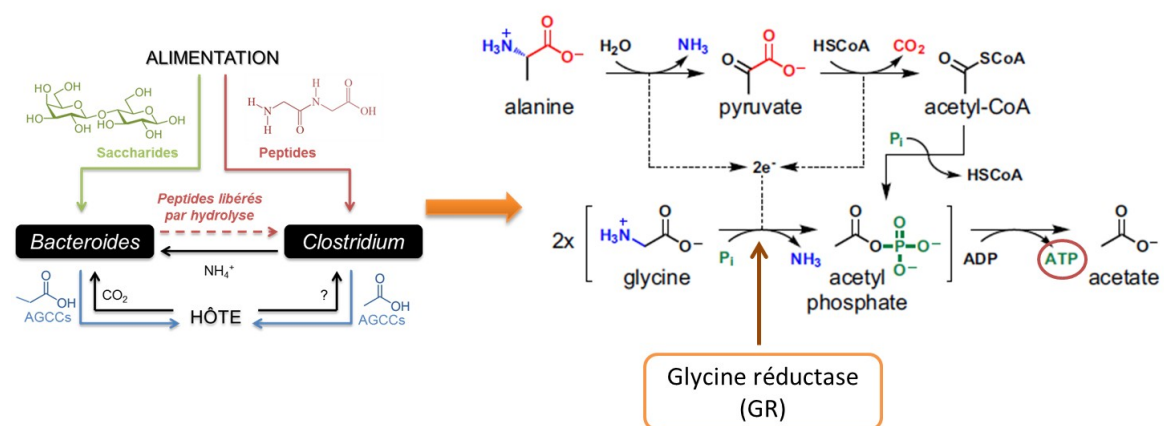


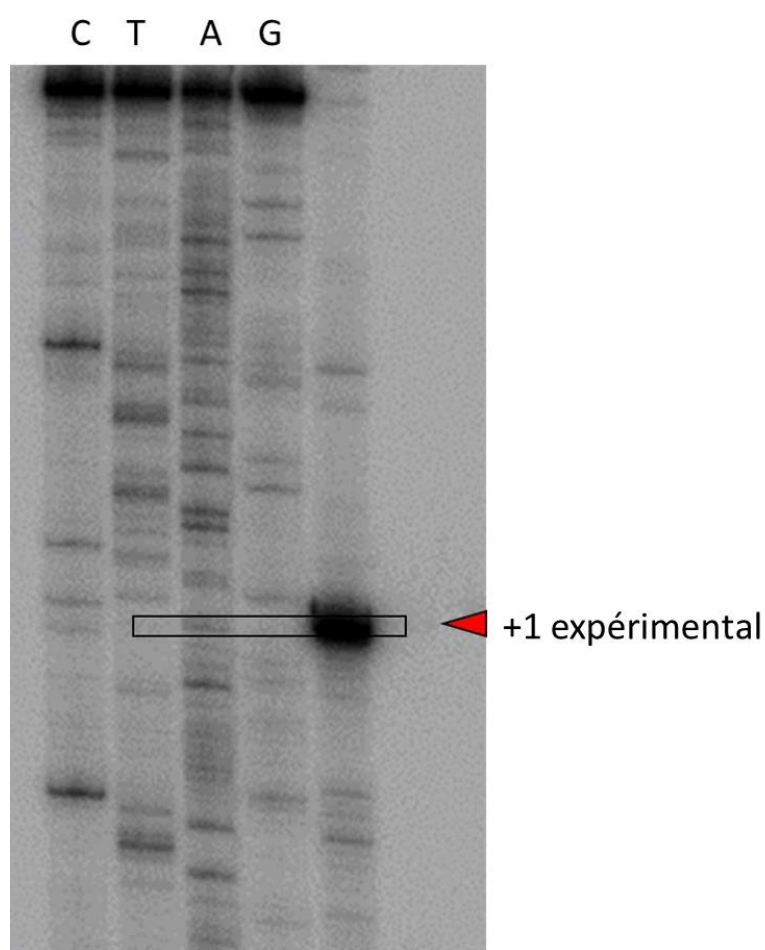
Figure 32. Schéma simplifié de la réaction de Stickland.

L'hypothèse a été émise que si la synthèse de la glycine réductase était inhibée, la bactérie serait privée d'une source d'ATP et pourrait être de ce fait moins compétitive. C'est cette hypothèse qui a soulevé l'intérêt pour ce riborégulateur et qui a été testée par la suite.

## VIII.2. Résultats

### VIII.2.1. RR-grdX, un riborégulateur structurellement inhabituel

Si la séquence globale du riborégulateur a été déterminée par bio-informatique, il manquait comme pour le symporteur, l'identification précise du site d'initiation de la transcription. Cette information a pu être donnée par l'utilisation de l'extension d'amorces à partir de la matrice d'ADN contenant tout le riborégulateur et remontant 103 pb en amont du début supposé du riborégulateur (Fig. 33).

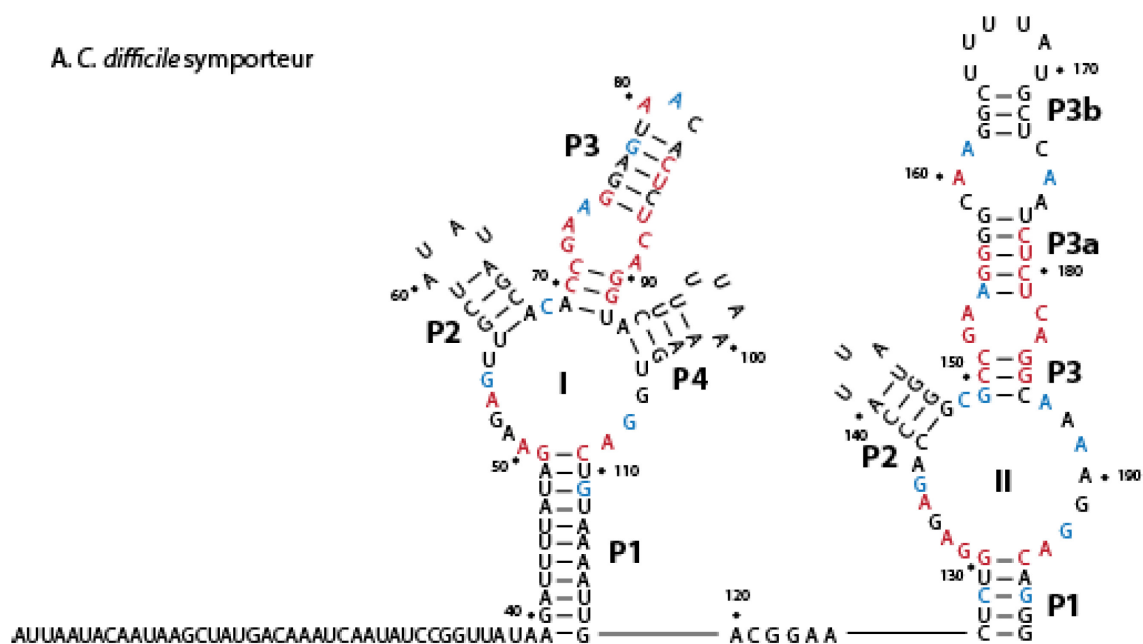


```
acaagaaatatgcaatatcaacatttggaaaaaaaaagtaaaatgtatacaaaatcacatataaaaaaactattgcaa  
atttatagaatgtttatatacttaagtcaacagagatagaaaaagctatttatagaaaacaaaggataaattgctta  
attttaaattttctgaaaaaaataataatccggatgaagagtatgggagactgctttataaaaagcgaccgaaggga  
caagtttaatcttacttaccbaaaagattagatgaaattctcaggtaaaggtgactgtacttggacggggctctgaaaa  
tgattttaacagggagagacaattgggagattt
```

Figure 33. Extension d'amorce pour l'ARN du riborégulateur *grdX*.

La séquence en vert représente l'amorce sens de la matrice tandis que la séquence soulignée indique l'amorce anti-sens de la matrice. La séquence théorique des aptamères est représentée en jaune tandis que la sonde est en gras.

A. *C. difficile* symporteur



B. *C. difficile* grdX

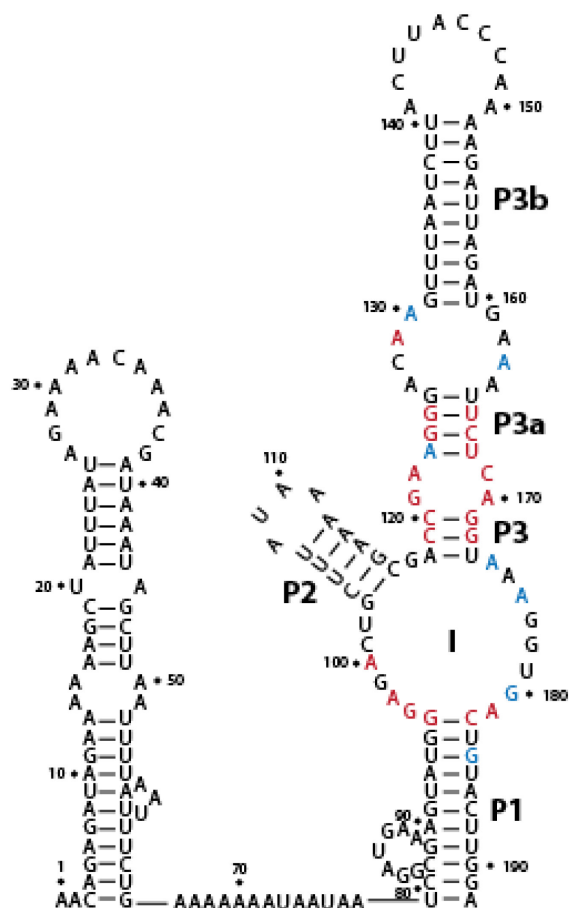


Figure 34. Comparaison des structures primaire et secondaires entre le RR-symp et le RR-grdX de *C. difficile*. Les nucléotides en rouge représentent une conservation >95% tandis que ceux en bleu >75%.

Cette technique a permis de déterminer que le début de la transcription s'effectuait un nucléotide en amont de ce qui avait été prédit. Ceci a mené à la détermination de la structure secondaire des aptamères du riborégulateurs à l'aide du programme en ligne mFold. Il a pu être ainsi constaté qu'en effet, le RR-grdX était un riborégulateur simple ne disposant que d'un seul aptamère. Cet aptamère est très semblable au second du tandem retrouvé chez le RR-symp (Fig. 34) et à fortiori à celui de *V. cholerae*. Le point le plus intéressant toutefois est la présence d'une longue séquence en 5' de l'aptamère qui forme une structure secondaire en une longue tige-boucle (l'ensemble tige-boucle et aptamère est donné en jaune sur la séquence de la Fig. 33).

#### VIII.2.2. Une fonctionnalité renouant avec la normale

Étant donné que les essais rapporteurs utilisant la  $\beta$ -glucuronidase n'étaient pas concluants chez *C. difficile*, ils n'ont pas été tentés pour le RR-grdX, même s'il serait intéressant de les faire. Les essais ont directement été conduits chez *B. subtilis* 168. Pour cela, le RR-grdX a été inséré en amont du gène codant pour la  $\beta$ -galactosidase (*lacZ*). Des essais préliminaires concluants avaient été entrepris en milieu riche (LB) mais la composition de ce milieu pouvant difficilement être connue avec certitude, la concentration en glycine présente en son sein bien que certainement faible gênait les expérimentations. Les essais en milieu pauvre ont été davantage concluants et ont permis d'observer une augmentation de l'activité enzymatique suite à l'ajout d'une concentration finale de 10 mM de glycine dans le milieu. Ainsi, une activation de l'expression d'un rapport 2,4 a pu être constatée en présence de ligand (Fig. 32).

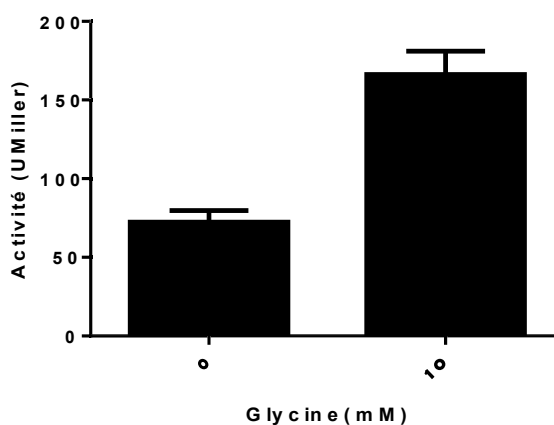


Figure 35. Activité galactosidase du RR-grdX chez *B. subtilis* N=3, les barres d'erreur représentent l'écart-type SD.



Suite à cela, et en accord avec la littérature, un mutant de liaison de la glycine a été construit. Cette mutation intervient au niveau de l'uracile 168 où interagit la glycine et est remplacé par une adénine. Comme ce qui était escompté, la mutation U168A entraîne un arrêt de la régulation du gène par le RR-grdX et l'expression du gène *lacZ* est toujours réprimée (Fig. 36).

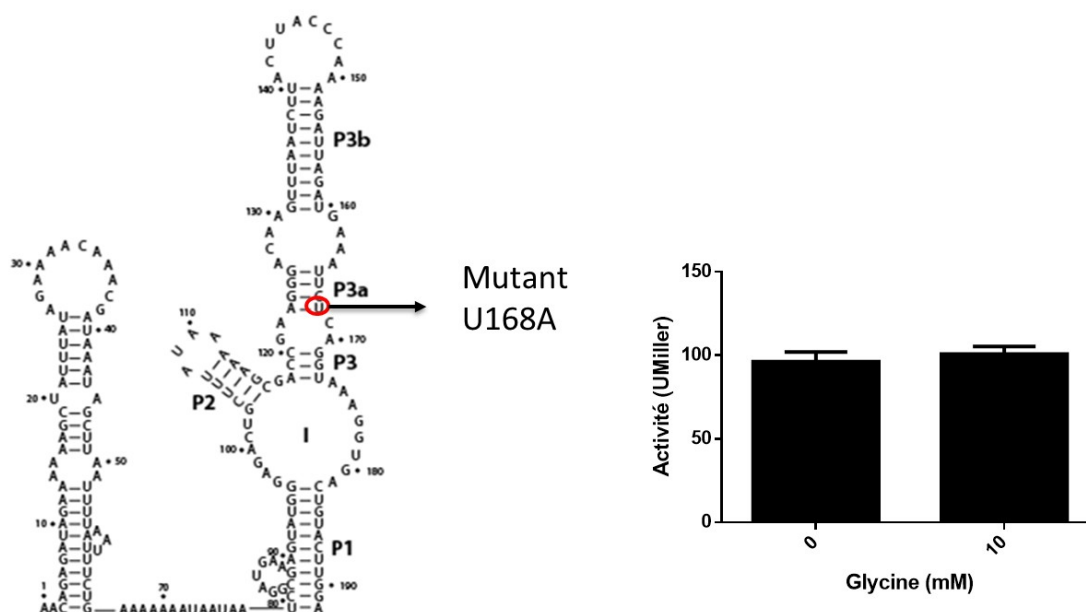


Figure 36. Activité galactosidase du mutant U168A de liaison du RR-grdX chez *B. subtilis* N=3, les barres d'erreur représentent l'écart-type SD.

### VIII.2.3. Un singlet aussi efficace qu'un tandem

Les essais p-galactosidase ont permis de démontrer que le RR-grdX était bien fonctionnel et activait la transcription du gène placé en aval. Néanmoins, son affinité pour la glycine restait à déterminer et pour cela la transcription *in vitro* à cycle unique a été employée. Cette technique permet de visualiser la transcription en fonction de la concentration en ligand, et ce en se limitant à une seule transcription par matrice d'ADN grâce à l'ajout d'héparine qui va venir se lier à l'ARN polymérase libre.

Tout d'abord, ne disposant que de la polymérase d'*E. coli* au laboratoire, le promoteur LacUV5 permettant la reconnaissance de l'ADN par cette enzyme a dû être ajouté en 5' du riborégulateur. Le RR-grdX a été ensuite testé avec différents tampons de transcription contenant différents sels (NaCl ou KCl) et différentes concentrations de rNTPs afin d'optimiser la réaction. Finalement ce sont les conditions avec le tampon au KCl et une concentration en rNTPs de 25μM qui ont été retenues (Fig. 37).

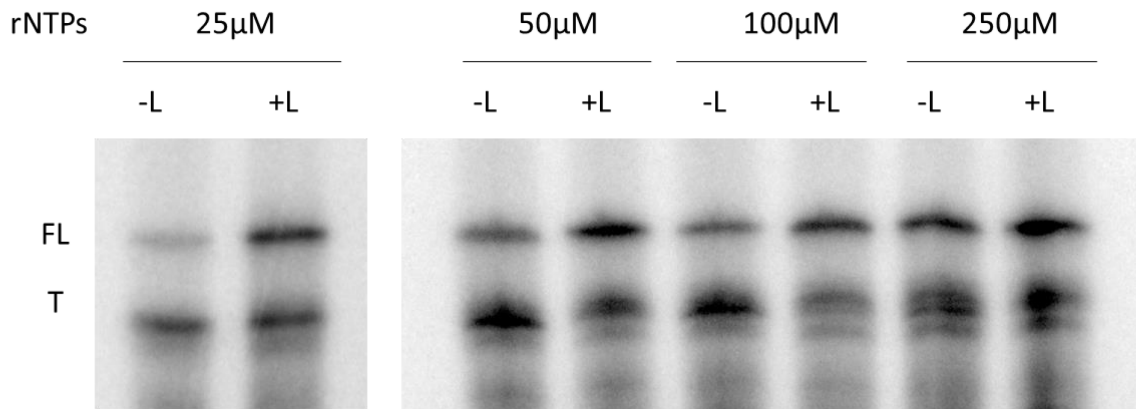


Figure 37. Tests pour déterminer la meilleure concentration en rNTPs.

Ces tests préliminaires montraient en outre bien le caractère activateur de la transcription, l'ajout de glycine (Ligand L sur la Fig. 37) augmente la proportion d'ARN transcrits complètement (FL pour *Full Length*) par rapport à ceux terminés prématurément (T pour *Terminated*). Un gradient de concentration en glycine de 10 nM à 10 mM a été réalisé (Fig. 38). La proportion de transcrits terminés tend à disparaître avec l'augmentation de la concentration en ligand tandis que celle de transcrits complets augmente. Une fois le ratio calculé et converti en pourcentage, il est possible de tracer un graphe et de déterminer le T50 à partir de celui-ci. Pour une construction *wild type* donc de phénotype sauvage pour la séquence du RR-grdX, le T50 obtenu expérimentalement est de 18,53 μM (Fig. 38).

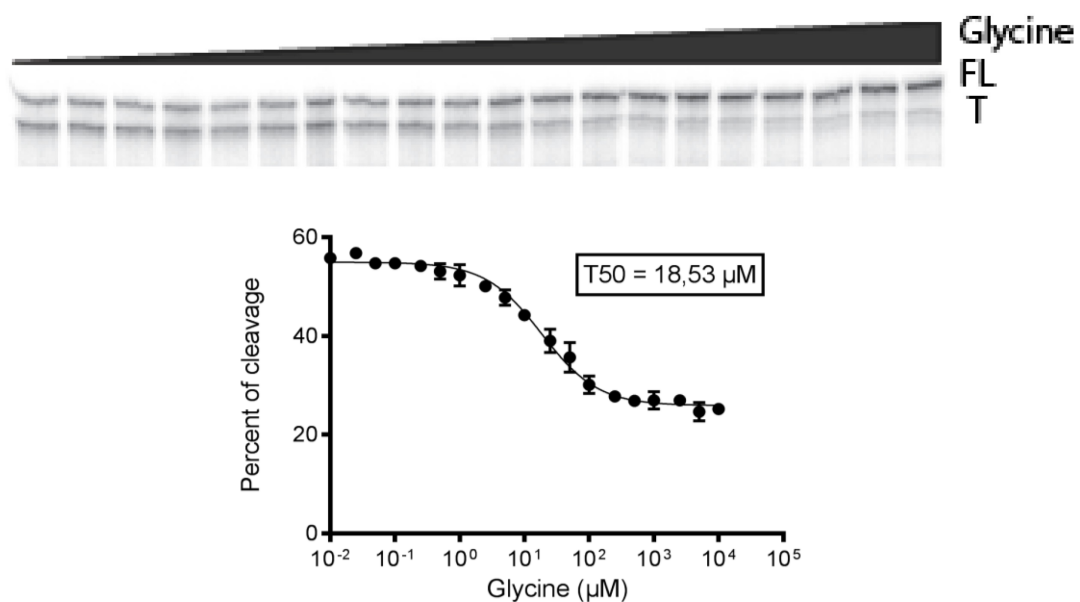


Figure 38. Gradient de concentration de glycine et détermination du T50 pour le RR-grdX.

Une fois ces données obtenues, il a été entrepris de déterminer si une telle tendance pouvait être observée lorsque le site de liaison du ligand était muté. La mutation utilisée, U168A, est la même que celle précédemment utilisée pour les essais rapporteurs chez *B. subtilis*. Il s'avère que la régulation est perdue (Fig. 39), il n'y a pas de modification de ratio entre le transcrit FL et le terminé.

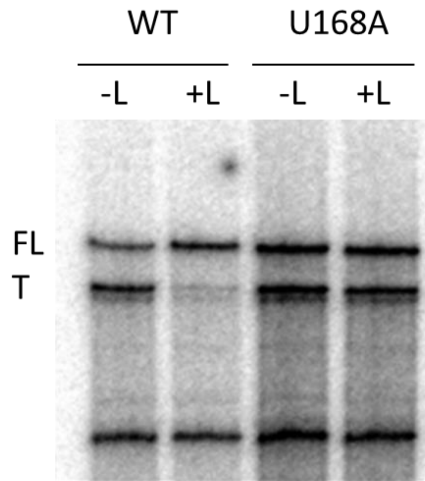


Figure 39. Transcription in vitro de la souche sauvage (WT) et du mutant de liaison (U168A) en absence (-L) ou en présence de glycine (+L).

Si la mutation du site de liaison du ligand abolit toute régulation, la question de l'utilité de l'aptamère fantôme demeurerait. En se basant sur les séquences impliquées dans le repliement des deux aptamères sur un riborégulateur classique ainsi que sur celles identifiées sur d'autres singlets, les nucléotides A175 et A176 ont été mutés en deux C (Annexe 5). À première vue, la régulation semblait perdue elle aussi. Mais un gradient de glycine a permis de déterminer que l'affinité du riborégulateur pour son ligand avait diminué d'un facteur 7, atteignant 128  $\mu$ M au lieu des 18,5  $\mu$ M pour une séquence WT (Fig. 41). Ici, on n'observe pas une diminution de la fraction des transcrit terminés mais une augmentation des transcrits pleine longueur.

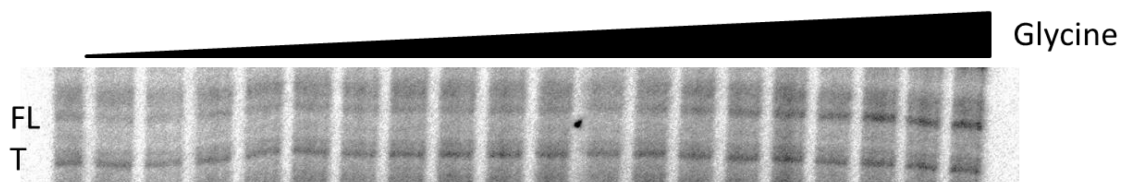


Figure 40. Transcription in vitro pour le mutant de repliement en fonction d'une concentration croissante de glycine.

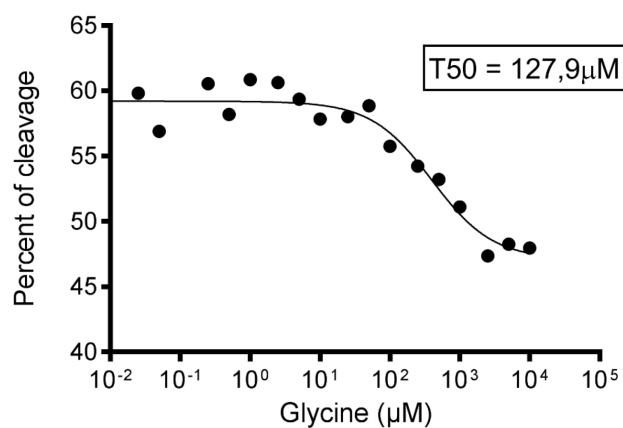


Figure 41. Détermination du T50 pour le mutant de repliement du RR-grdX.

#### VIII.2.4. Le RR-grdX, un candidat à écarter

De même que pour le symporteur, une fois la fonctionnalité du riborégulateur démontrée, il a fallu déterminer si l'inactivation du gène était délétère pour la bactérie. Le mutant inactivé par ClosTron a été comparé à la souche sauvage en milieu riche (TY, Fig. 42) et en milieu pauvre (Fig. 43).

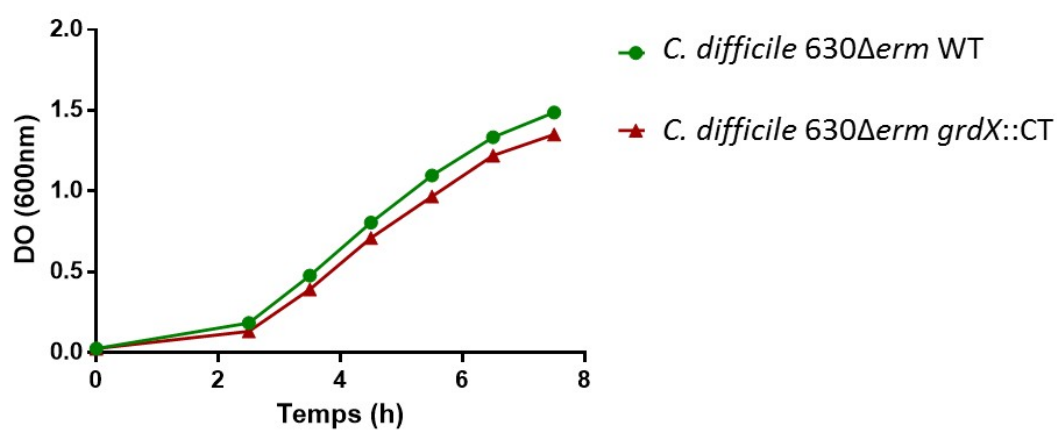


Figure 42. Courbes de croissance de la souche sauvage (WT) et du mutant inactivé de *grdX* en milieu riche. N=3.

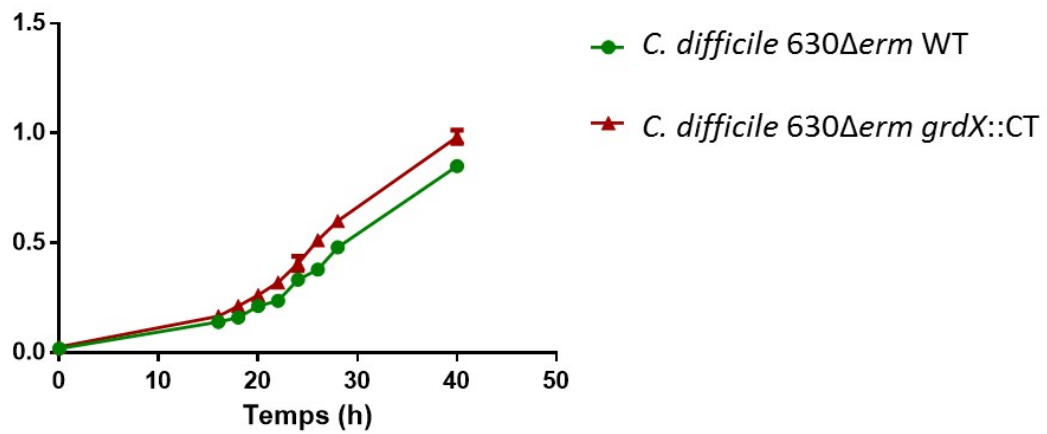


Figure 43. Courbes de croissance de la souche sauvage et du mutant inactivé de *grdX* en milieu pauvre. N=3.

Il n'y pas de différence significative observée en termes de croissance entre les deux souches que ce soit en milieu riche ou en milieu pauvre. Le mutant a tendance, comme pour le symporteur, à croître un peu plus rapidement que la souche sauvage mais c'est sans doute également dû à une plus faible charge métabolique (Fig. 43).

## IX. Discussion

Les objectifs de ce mémoire étaient de parvenir à une meilleure compréhension des riborégulateurs glycine retrouvés chez *C. difficile* afin d'être en mesure d'évaluer s'ils pourraient être utilisés comme cible thérapeutique. En effet, cette bactérie pathogène est de plus en plus résistante aux antibiotiques actuels et représente une large charge pour le système de santé que ce soit sur un plan humain ou financier. Les expériences réalisées au cours de cette maîtrise ont pu donner certains éclaircissements sur la structure, le fonctionnement ou l'impact de ces structures d'ARN.

### Organisation structurale des riborégulateurs glycine

Pour le RR-symp, l'organisation des aptamères en tandem est typique des riborégulateurs glycine retrouvés chez d'autres organismes tels *V. cholerae* ou *B. subtilis*. Néanmoins l'identification d'une longue séquence de 38 nt en amont du premier aptamère par l'extension d'amorce suggère la présence d'une petite hélice P0 (UCCG), aussi appelée *Kink-turn*, qui se lierait au niveau de la séquence entre les aptamères (CGGA) (Annexe 4). Cette liaison supplémentaire permettrait de stabiliser l'ensemble et d'en permettre une compaction maximale voire même d'augmenter l'affinité du riborégulateur pour le ligand en présence de quantités plus faibles de magnésium (Baird and Ferré-D'Amaré, 2013). Ce type d'hélice a déjà été montré dans des riborégulateurs glycine et est largement conservée (>90%) chez *B. subtilis*, *Fusobacterium nucleatum* et *V. cholerae* notamment (Kladwang et al., 2012; Sherman et al., 2012).

Pour le RR-grdX, les structures secondaires ont permis de mettre en évidence une structure supplémentaire de 78 nt en amont de l'aptamère. Il s'agit d'une longue tige-boucle regroupant 3 boucles. L'année dernière, 17 riborégulateurs putatifs montrant de telles structures ont été identifiées par alignement de séquences pour d'autres riborégulateurs glycine chez d'autres organismes (Ruff et al., 2016). Ces recherches s'étaient d'abord effectuées pour identifier de nouveaux riborégulateurs en tandem ce qui illustre mieux la conservation des séquences et ce même lorsque la moitié est manquante. Les auteurs ont également montré que bien qu'un seul aptamère était fonctionnel, le riborégulateur faisait preuve d'affinités comparables à leurs homologues en tandem. Les résultats obtenus ici ont aussi permis d'appuyer cette observation tout en confirmant

l'utilité de cet aptamère fantôme qui pourrait potentiellement être un vestige d'un aptamère autrefois fonctionnel. Si ce fantôme n'est plus en mesure de lier un ligand, il reste important pour le repliement du riborégulateur. En effet, nous avons pu voir qu'en mutant le site de l'interaction  $\alpha$ , la capacité du riborégulateur à activer la transcription était largement diminuée. Les sites d'interactions étant largement conservés même dans ces vestiges tandis que le reste de la structure peut varier considérablement. Il en va de même lorsque l'aptamère fantôme se situe en 3' (singlet de type 1) de l'aptamère fonctionnel bien que ce soit l'interaction  $\beta$  qui soit impactée. Étonnamment, bien que les singlets soient présentés comme possédant également une séquence *kink-turn*, ce n'est pas le cas pour le RR-grdX.

D'autre part, les singlets sont retrouvés en amont des mêmes types de gènes que les riborégulateurs en tandem. Un tiers d'entre eux sont ainsi retrouvés en amont de gènes impliqués dans le clivage de la glycine pour les singlets de type 2 (cas qui nous intéresse ici) tandis que cette proportion monte jusqu'à environ 85% pour les singlets de type 1.

#### Essais rapporteurs et fonctionnalité *in vivo*

Il est toujours plus intéressant de procéder aux essais rapporteurs dans l'organisme ciblé mais malheureusement ici cela n'a pas été possible. De multiples essais en changeant les conditions de lyse des cellules ont été tentés comme l'utilisation de l'endolysine ou encore de la sonication. La perméabilisation des cellules bien que prometteuse n'a pas permis d'obtenir des résultats plus probants. L'hypothèse que le temps en présence de ligand une fois les cellules transférées dans du milieu minimum n'était pas suffisant a également été émise. Pour déterminer si ce facteur temps était la raison de l'impossibilité d'obtenir un effet ligand, les bactéries ont été inoculées à 5% dans le milieu minimum avec ou sans ligand au lieu d'être simplement resuspendues. Elles ont ensuite été mises en culture jusqu'à obtenir une DO de 0,7 ce qui pouvait prendre jusqu'à 36h. Malheureusement ceci n'a pas permis d'obtenir l'effet escompté. Une hypothèse serait que la glycine peine à se rendre au sein des cellules, une autre serait que ces essais impliquaient l'utilisation d'une copie plasmidique du gène rapporteur et donc potentiellement de plusieurs copies du riborégulateur et du gène au sein de la bactérie ce qui pourrait impacter l'effet ligand. Cependant, la raison pour laquelle les essais  $\beta$ -glucuronidase chez *C. difficile* n'ont pas fonctionné restent non élucidées. Le nom de cet

organisme nommé ainsi en 1935 en raison des difficultés à l'étudier est toujours d'actualité, que ce soit avec des essais rapporteurs infructueux ou une impossibilité à transformer l'organisme qui nécessite l'utilisation de ClosTrons et d'organismes intermédiaires. En effet, les essais de mutagenèse directe par recombinaison homologue n'ont produit que des intégrants instables (Heap et al., 2007) bien que d'autres méthodes aient montré des résultats positifs (O'Connor et al., 2006).

La décision fut par la suite prise d'effectuer les essais de gènes rapporteurs chez *B. subtilis* qui est un organisme Gram(+) comme *C. difficile* mais qui est bien plus facile à travailler notamment car aérobie. Pour essayer de préserver au maximum les conditions originelles du riborégulateur le promoteur constitutif du gène de *C. difficile* a été utilisé au lieu de prendre celui du plasmide comme c'était le cas avec le plasmide pRPF144 chez *C. difficile*. Néanmoins ce plasmide ayant déjà fait ses preuves dans la littérature pour des essais  $\beta$ -glucuronidase, ceci n'est pas censé être la cause de l'impossibilité d'avoir un effet ligand (Fagan and Fairweather, 2011). Il faut toutefois nuancer hypothèse que les essais entrepris par Fagan et Fairweather ne faisaient pas intervenir de riborégulateurs. Il a toutefois été jugé plus sage de conserver le promoteur constitutif. Le fait que le plasmide utilisé chez *B. subtilis* (pDG1661) ait une insertion ectopique dans le génome pourrait peut-être avoir un impact car il n'y a pas besoin de maintenir une pression de sélection par l'ajout d'antibiotique dans le milieu et limiterait à une seule copie le gène rapporteur. On peut supposer que la charge de devoir répliquer un plasmide ainsi que le nombre de copies du gène et du riborégulateur aient pu avoir un impact sur l'activité.

Les essais rapporteurs chez *B. subtilis* ont ainsi permis de confirmer que le RR-symp était bien un riborégulateur glycine répresseur de la transcription comme ce qui était pensé d'après les données de RNAseq. Ceci tranche avec les riborégulateurs glycine habituels qui sont davantage des activateurs de la transcription. Dernièrement il a été suggéré que des singlets de type 2 pourraient être des inhibiteurs de la transcription (Ruff et al., 2016). De même les essais pour le RR-grdX ont permis de confirmer le caractère activateur du riborégulateur. La mutation du site de liaison du ligand chez ce dernier a permis également d'abolir tout effet du riborégulateur sur l'expression du gène comme ce qui était attendu. Des essais semblables chez le RR-symp seraient intéressants à mener.



## Détermination de l'affinité des riborégulateurs glycine

Les analyses *in vitro* avaient pour but de déterminer l'affinité des riborégulateurs pour la glycine mais ont malheureusement eu un succès mitigé. En effet, si la cartographie a permis dans le cas du second aptamère de RR-symp de recueillir des données, ce n'est pas le cas de la transcription *in vitro* à cycle unique où aucun effet ligand n'a pu être constaté. De même une analyse par SHAPE (*Selective 2'-hydroxyl acylation analyzed by primer extension*) qui consiste également à étudier la protection de certains nucléotides aux attaques d'un réactif, n'a pas donné de résultats exploitables car là encore aucun effet ligand n'a pu être observé. C'est d'ailleurs l'absence de résultats probants malgré les tentatives qui a mené à la mise de côté de ce riborégulateur et explique une étude moins étendue de son fonctionnement. Cependant, les zones de protection identifiées à l'aide de la cartographie chimique (*In line probing*) correspondent à celles identifiées par l'équipe de Mandal et Breaker en 2004 lors de la découverte des riborégulateurs glycine (Mandal et al., 2004). De plus les  $K_d$  déterminés de l'ordre de  $4\mu\text{M}$  sur le second aptamère concordent avec les données obtenues par l'équipe de Ruff sur *V. cholerae* qui oscillent entre 2 et  $4\mu\text{M}$  (Ruff and Strobel, 2014). Il a également été tenté pour faire fonctionner les essais de transcription *in vitro* à cycle unique de déplacer l'initiation de la transcription. En effet, l'amorce sens est constituée du promoteur LacUV5 d'*E. coli* associé au début de la séquence à transcrire. Le choix de cette séquence s'effectue en fonction des nucléotides puisque seuls trois doivent être présents pour permettre le blocage de l'ARN polymérase avant d'ajouter le mélange de redémarrage. Trouver de telles séquences suffisamment longues ( $> 9\text{nt}$ ) peut être difficile et plusieurs positions ont été testées notamment avec ou sans la séquence en 5' impliquée dans le *kink-turn*. Cela n'a malheureusement pas permis l'observation d'effet ligand sur la transcription.

Pour le RR-grdX, les expériences de transcription *in vitro* ont permis de confirmer que la régulation activait l'expression du gène et que celle-ci intervenait à un niveau transcriptionnel. La concentration minimale nécessaire pour obtenir 50% de terminaison (le T50) observée est de  $18,53\mu\text{M}$ . Elle est du même ordre de grandeur bien que plus élevée que le  $K_d$  retrouvé dans la littérature pour les riborégulateurs glycine de *V. cholerae* (Ruff and Strobel, 2014). En effet, le  $K_d$  identifie l'affinité entre le ligand et le riborégulateur mais cela ne signifie pas que celle-ci est suffisante pour induire un changement de conformation de l'ARN et donc une régulation. Une valeur de T50 peut donc être égale au  $K_d$  dans de rares cas mais sera généralement plus élevée. De même,

lorsqu'ils ont cherché à muter l'interaction  $\alpha$  responsable du repliement, le  $K_d$  a augmenté d'un facteur 7,7 ce qui correspond à l'augmentation du T50 (de 7 fois) qui a pu être observée expérimentalement sur le riborégulateur *grdX*. Comme pour les essais rapporteurs, la mutation du site de liaison a mené vers une perte de la régulation comme c'était attendu (Ruff and Strobel, 2014).

D'autres essais *in vitro* tels que de la cartographie chimique ont été entrepris sur ce riborégulateur (RR-*grdX*) sans que des résultats concluants aient pu être obtenus. En effet, aucune modification significative suite à l'ajout du ligand n'a pas été constatée. Une hypothèse pouvant expliquer ceci serait que les structures déjà présentes en absence de ligand seraient de même nature que celles induites par la liaison de la glycine. De ce fait, les mêmes nucléotides pourraient rester accessibles et aucun changement ne serait alors détectable. Des données obtenues au laboratoire Lafontaine sur le riborégulateur magnésium *mgtA* de *E. coli* pourtant extrêmement semblable en termes de séquences et de structures à celui de *Salmonella enterica* montrent qu'il n'est pas régulé de la même manière. Il est donc théoriquement possible que là où un effet avait été constaté chez *B. subtilis* on n'en observe pas de similaire chez *C. difficile*. Les conditions expérimentales auraient sans doute pu être optimisées mais il ne nous a pas semblé pertinent de passer plus de temps sur ces expériences étant donné les informations apportées par la transcription *in vitro* à cycle unique.

#### Courbes de croissance et inactivation de gènes

Les courbes de croissance en milieu pauvre peuvent sembler intéressantes pour le RR-symp du fait du retard de croissance constaté et pourraient indiquer un effet délétère de l'inactivation du gène. Néanmoins celui-ci n'est qu'un retard et le mutant rattrape par la suite la souche sauvage au niveau de la phase stationnaire. Une hypothèse expliquant ce phénomène serait que la bactérie ait besoin de temps pour actionner un mécanisme accessoire qui lui permettrait de compenser la perte de son symporteur. Une fois ce mécanisme en place, elle serait parfaitement en mesure de croître comme la souche sauvage.

Les expériences réalisées sur le riborégulateur de *grdX* ont permis de déterminer que la bactérie n'a pas besoin de ce gène pour être viable. Il est intéressant de comparer ces résultats avec ceux obtenus par Bouillaut *et al.* qui ont muté par ClosTron eux aussi le gène *grdA* présent en aval de *grdX* (Bouillaut *et al.*, 2013). Ce gène, tout comme *grdB*, encode une sous-unité contenant de la sélénocystéine appartenant à la glycine réductase. Ils ont pu montrer que cet opéron *grd* était également sous le contrôle de la proline. En effet, l'ajout de proline dans le milieu permettait d'augmenter la croissance du mutant *grdA*, ce qui suggère là aussi que la glycine réductase n'est pas essentielle à la croissance bactérienne. Même si ce résultat était connu, nous espérions qu'en mutant une autre sous-unité potentielle de la glycine réductase (composée des sous-unités *grdA*, *grdB* et *grdC*) nous obtiendrions des effets néfastes pour la bactérie. Nous aurions pu observer des effets polaires ou, puisque ce gène est sous le contrôle d'un riborégulateur et de ce fait potentiellement important à réguler, un effet sur une autre voie métabolique. Cependant, il s'avère qu'il n'en est rien. En effet, même si contrairement au symporteur il s'agissait potentiellement d'une activation de la transcription, les courbes de croissance ont permis de constater que la présence ou non du gène et donc de la glycine réductase n'avait pas d'impact sur la croissance de la bactérie que ce soit en milieu riche ou en milieu pauvre, plus proche des conditions rencontrées dans le tractus digestif. Le milieu pauvre représente en effet ce que pourrait être l'intestin non pas en raison d'une faible quantité de nutriments présente mais en raison de la grande compétitivité retrouvée pour les nutriments due à une flore microbienne très dense.

Si le blocage du riborégulateur, et donc du gène, s'était révélé probant, des tests dans un modèle murin auraient été entrepris. En effet, il a été montré précédemment dans le laboratoire avec le riborégulateur guanine chez *C. difficile* qu'un retard significatif de croissance en milieu minimum pouvait mener à une perte de compétitivité chez la bactérie. Des essais de co-infection chez la souris où des spores de souche sauvage ainsi que des spores de souches mutées ont été administrés par gavage puis les fèces des souris ont été récupérés durant plusieurs jours puis mis en culture pour le dénombrement sur gélose des différentes souches de *C. difficile*. Cette discrimination est rendue possible grâce au ClosTron qui a inséré un gène de résistance à l'érythromycine, il suffit alors de soustraire les colonies mutantes au nombre total de colonies pour déterminer les bactéries sauvages. Ces essais ont montré que le mutant de *guaA*, gène codant pour une GMP

synthase, n'était pas en mesure de se maintenir dans le tractus digestif de l'animal ce qui fait du riborégulateur *guaA*, une cible thérapeutique à explorer.

Les essais chez le modèle murin n'ont toutefois pas été entrepris car il a été depuis montré au laboratoire Fortier qu'un retard de croissance n'était pas forcément synonyme d'effet *in vivo*. En effet, une croissance très diminuée pour un mutant du gène *nadA* avait été observée et les tests de co-infection ont été réalisés. Il s'avère qu'il n'y a pas eu d'impact sur la croissance de la souche dans le tractus digestif de la souris. Le retard de croissance observé chez le mutant du gène du symporteur glycine n'a donc pas été jugé suffisant pour entreprendre de tels essais qui sont coûteux et relativement longs. Néanmoins il pourrait être intéressant de procéder à des essais de co-culture *in vitro* en inoculant le mutant et la souche sauvage pour faire une première analyse de compétitivité et déterminer si des essais chez la souris pourraient quand même être utiles. Si des résultats probants sont obtenus chez le modèle murin, ils permettent d'envisager le design d'une molécule analogue du ligand mais non métabolisable. Cet analogue pourrait venir se fixer au riborégulateur et ainsi avoir des effets antibiotiques. Bien évidemment ces essais requièrent de nombreux ajustements car si les mammifères ne possèdent pas de riborégulateurs dans les connaissances actuelles, ce n'est pas le cas du reste du microbiote intestinal ce qui pourrait entraîner des effets délétères pour le microbiote et *in fine* pour l'hôte en raison d'une dysbiose.

#### Mutations et fonctionnalité des riborégulateurs glycine

Il est intéressant de constater que l'aptamère fantôme, bien qu'incapable de lier le ligand, est toujours important pour le repliement du riborégulateur et participe donc à la fonctionnalité du riborégulateur. En effet, avant de procéder aux mutations de repliement telles que présentées précédemment, une tentative de mutation où l'aptamère fantôme était retiré au complet a été faite. Celle-ci n'a donné aucun résultat en transcription *in vitro*. Un tel effet était déjà suggéré dans la publication de Ruff *et al.* quand ils ont étudiés les singlets (Ruff et al., 2016). S'ils ont réalisé des mutants sur les singlets de type 1 (*L. monocytogenes*) en enlevant complètement l'aptamère fantôme qui se trouvait en 3', ce n'est pas le cas lorsque les singlets de type 2 (*D. hafniens*) ont été étudiés. Leurs expériences sur le repliement n'impliquaient que des aptamères mutés à des sites précis pour empêcher le repliement plutôt que leur suppression. Cela semble confirmer

l'hypothèse que la liaison du ligand stabilise la tige de l'aptamère et défavorise la formation d'une tige-boucle alternative qui serait à l'origine du changement de conformation. L'aptamère fantôme, d'une structure plus longue chez les type-2, permettrait donc de s'associer à l'aptamère de façon plus stable ou de former directement l'interaction de type  $\alpha$  (Ruff et al., 2016).

## X. Conclusions

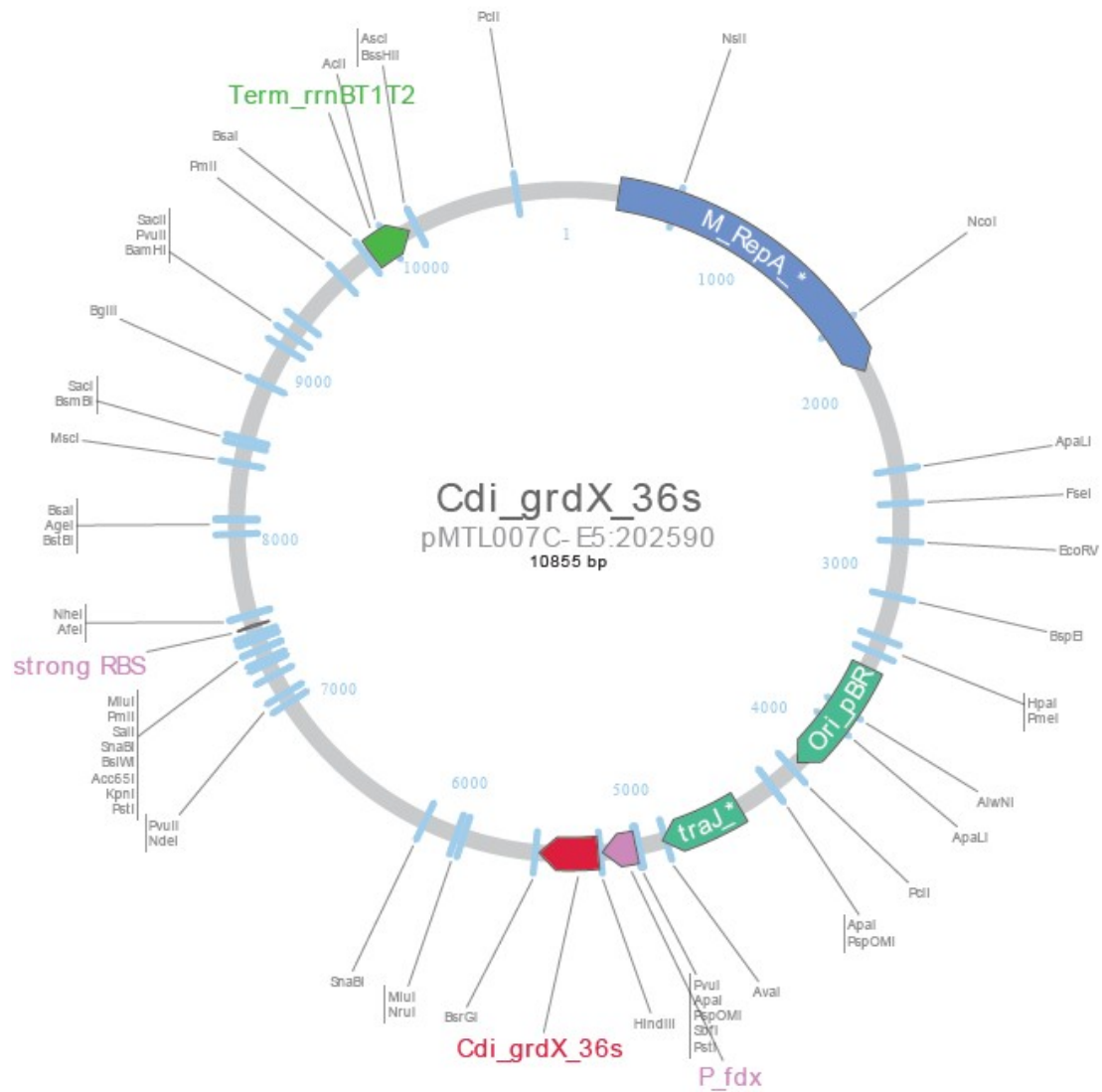
L'étude des riborégulateurs glycine chez *C. difficile* semblait prometteuse du fait de l'importance de cet acide aminé dans le métabolisme bactérien. Ces structures n'avaient jamais été analysées chez cet organisme et d'un point de vue fondamental, les travaux réalisés durant cette maîtrise ont permis de mieux les comprendre. Les données *in vitro* obtenues corroborent celles reportées pour d'autres riborégulateurs de cette classe chez des organismes plus conventionnels tels que *B. subtilis* ou *V. cholerae*.

D'autres essais *in vitro* et *in vivo* sont envisagées pour ce projet comme des mutants de liaison (U87A et U181A) et de repliement  $\alpha$  pour le RR-symp en essais rapporteurs (Annexe 4) ainsi que des études plus poussées sur la plateforme d'expression des deux riborégulateurs. Si celle du RR-grdX semble assez simple (Annexe 5), il existe deux séquences terminatrices potentielles dans celle du RR-symp. Des essais rapporteurs sur des séquences tronquées en aval du premier terminateur permettrait de discriminer les deux. Néanmoins, le terminateur retrouvé le plus en 3' serait un meilleur candidat car retrouvé après la séquence de l'anti P1. De plus, la séquence plus courte du premier terminateur et la séquence poly-U la suivant étant interrompue par une adénine, il semble plus probable que ce soit un site de pause plutôt qu'une réelle terminaison de la transcription.

Si les résultats obtenus pendant cette maîtrise ne supportent pas l'hypothèse de départ et ne permettent pas d'envisager une application thérapeutique, ils ne sont pas pour autant dénués d'intérêt. En effet même si ce n'est que grâce à une petite pierre à l'édifice, ces résultats identifient et étudient pour la première fois les riborégulateurs glycine chez *C. difficile* et ajoutent à la connaissance fondamentale de cet organisme. Ne dit-on pas d'ailleurs que pour combattre efficacement un ennemi il faut bien le connaître ? Il se peut que ces résultats servent à jour d'une quelconque manière dans la lutte contre ce pathogène opportuniste aux effets dévastateurs tant humains que financiers dans nos sociétés.

## XI. Annexes

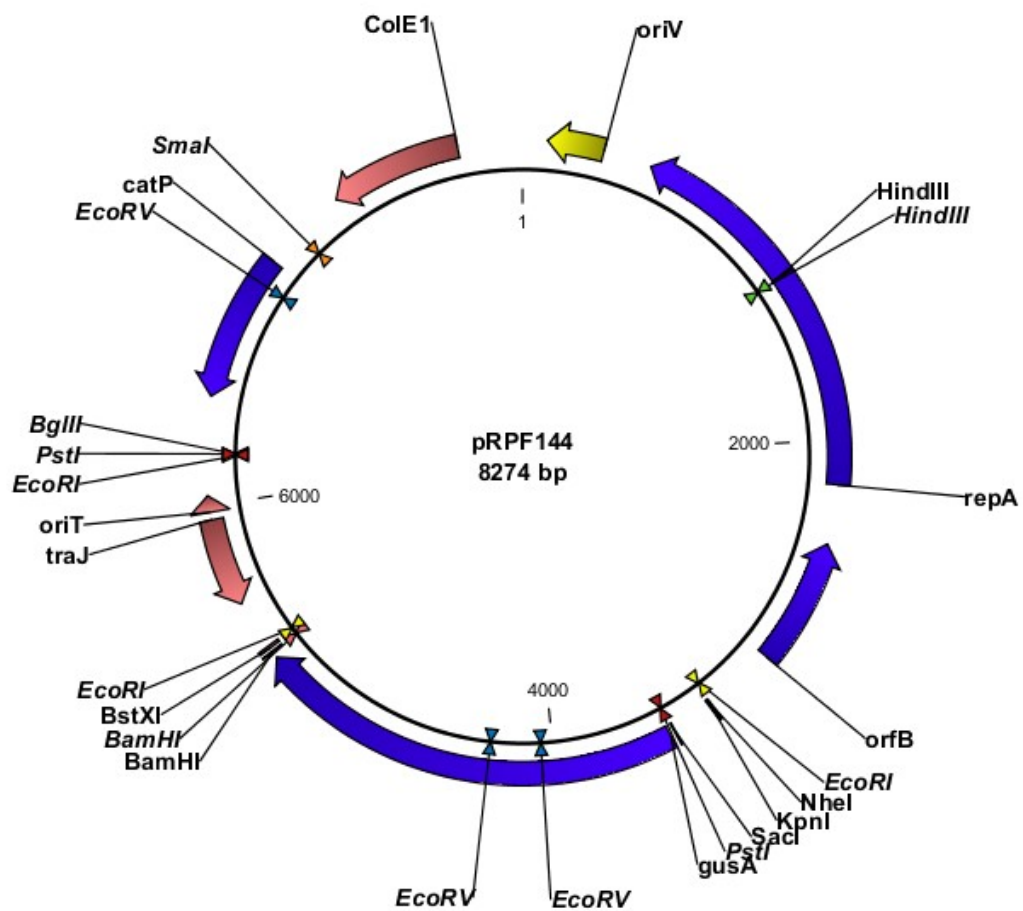
### Annexe 1. Carte du plasmide pMTL007C-E5 pour ClosTron.



### Légende des cadres de lecture ouverts (Heap et al., 2010)

RepA	région de réplcation
Ori_pBR	Origine de réplcation
TraJ	Fonction de transfert de la région RP4 de Ori_pBR
P_fdx	Terminateur
Strong RBS	Site de liaison du ribosome
Term_rrnBT1T2	Terminateur

Annexe 2. Carte du plasmide pRPF144.

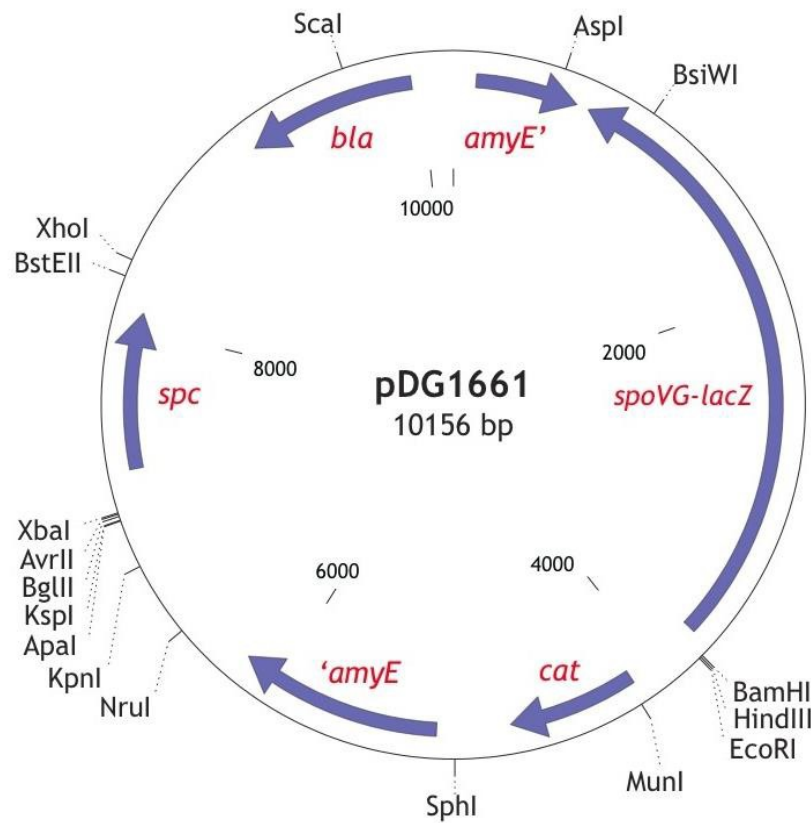


**Légende des cadres de lecture ouverts** (Fagan and Fairweather, 2011)

OriV	Origine de réplication
RepA	Région de réplication de <i>C. difficile</i>
OrfB	Région de réplication de <i>C. difficile</i>
gusA	Code pour la $\beta$ -glucuronidase avec le promoteur constitutif P <sub>cwp2</sub>
TraJ	Fonction de transfert de la région RP4 de l' <i>oriV</i>
catP	Code pour le gène de résistance de <i>C. perfringens</i> au chloramphénicol
ColE1	Région de réplication de <i>E. coli</i>



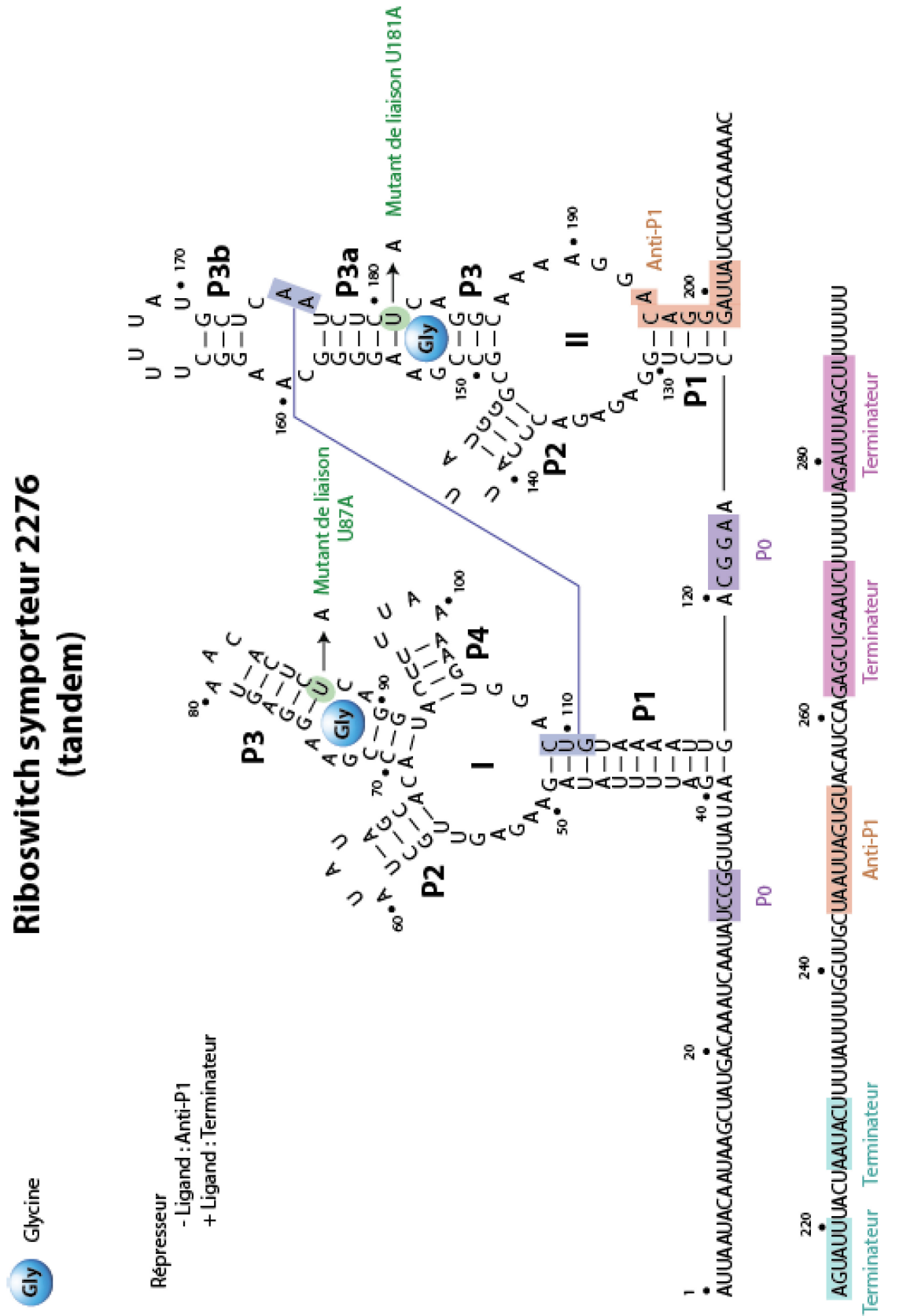
Annexe 3. Carte du plasmide pDG1661.



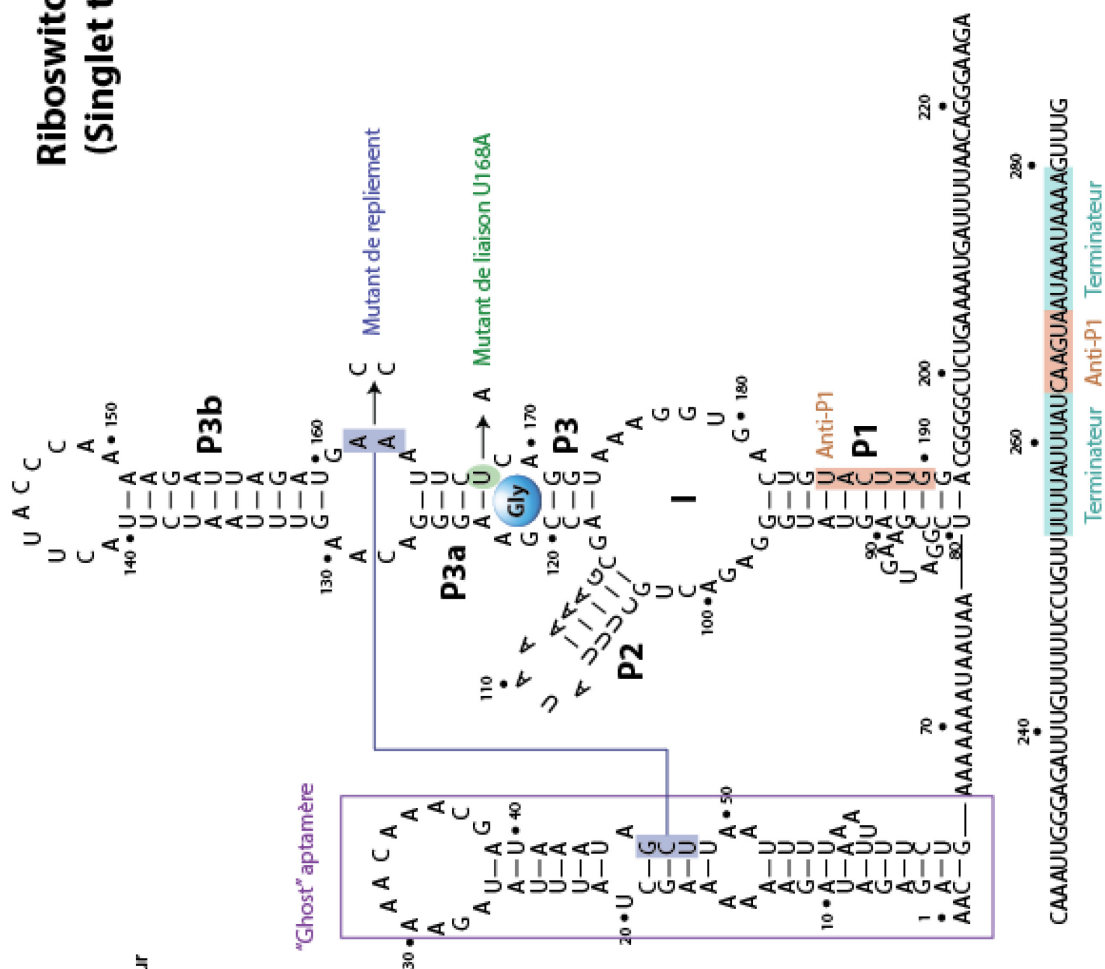
**Légende des cadres de lecture ouverts** (Guérout-Fleury et al., 1996)

<i>amyE'...amyE</i>	Segments 5' et 3' du gène <i>amyE</i> de <i>B. subtilis</i> 168
<i>spoVG-lacZ</i>	Séquence codante d' <i>E. coli</i> pour LacZ fusionnée au RBS <i>spoVG</i> de <i>B. subtilis</i>
<i>spc</i>	Code pour l'adényltransférase pour la spectinomycine pour sélection par antibiotique chez <i>E. coli</i> et <i>B. subtilis</i> (100µg/ mL)
<i>cat</i>	Code pour l'acétyl transférase pour le chloramphénicol pour sélection par antibiotique chez <i>E. coli</i> et <i>B. subtilis</i> (5µg/ mL)
<i>bla</i>	Code pour la β-lactamase pour sélection par antibiotique chez <i>E. coli</i> seulement (50µg/ mL)

**Riboswitch symporteur 2276  
(tandem)**



### Riboswitch grdX (Singlet type 2)



## Remerciements

Je tiens à remercier Louis-Charles Fortier et Daniel Lafontaine pour m'avoir accueillie au sein de leurs laboratoires respectifs. Ce fut deux années riches en enseignements et même si elles ont pu être difficiles parfois, je ne les regrette pas. Merci pour les conseils et le support mais également pour les BBQ, poutine's night, Oktoberfest et autres partys de Noël. Merci à mes conseillers Vincent Burrus et Alfredo Menendez pour leur temps et leurs conseils.

Je tiens également à remercier tous mes collègues avec des mentions spéciales pour Maicol Ospina-Bedoya et François Kirouac pour tous les *free hugs* qui m'ont bien remonté le moral quand j'en avais besoin mais également pour nos découvertes musicales dans le laboratoire et nos blagues pas drôles (je penserai à garder un couteau à beurre dans ma voiture, promis Mathieu Larocque !). Merci aussi pour les soirées jeux de société que ce soit avec un labo ou l'autre. Merci Mélodie Berthoumieux pour les soirées filles et les gros délires. Je ne sais pas par contre si je dois vous remercier Jean-Christophe Berger-Dancause, Jean-François Nadon et Jonathan Grondin pour la conversion à Magic the Gathering ... L'avenir et les cartes nous le diront !

Je remercie également Maxime Thomas Le Déoré pour le support moral depuis plus de 10 ans (et c'est pas fini mon grand !) et pour m'avoir aidé sur certaines figures de certaines de ce mémoire.

Un énorme merci à ma famille en France et en Irlande. Merci Maman de m'écouter me plaindre régulièrement et pour avoir toujours cru en moi.

## Liste des Références

- Alang, N., and Kelly, C.R. (2015). Weight Gain After Fecal Microbiota Transplantation. *Open Forum Infect. Dis.* 2, ofv004-ofv004.
- Ames, T.D., and Breaker, R.R. (2011). Bacterial aptamers that selectively bind glutamine. *RNA Biol.* 8, 82–89.
- Andrés-Lasheras, S., Bolea, R., Mainar-Jaime, R.C., Kuijper, E., Sevilla, E., Martín-Burriel, I., and Chirino-Trejo, M. (2016). Presence of *Clostridium difficile* in pig faecal samples and wild animal species associated with pig farms. *J. Appl. Microbiol.*
- Artsimovitch, I., and Henkin, T.M. (2009). In vitro approaches to analysis of transcription termination. *Methods* 47, 37–43.
- Astbury, W.T., and Bell, F.O. (1938). Some recent developments in the X-Ray study of proteins and related structures. *Cold Spring Harb. Symp. Quant. Biol.* 6, 109–121.
- Avery, O.T., Macleod, C.M., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: inducing of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J. Exp. Med.* 79, 137–158.
- Baird, N.J., and Ferré-D'Amaré, A.R. (2013). Modulation of quaternary structure and enhancement of ligand binding by the K-turn of tandem glycine riboswitches. *RNA* 19, 167–176.
- Baker, J.L., Sudarsan, N., Weinberg, Z., Roth, A., Stockbridge, R.B., and Breaker, R.R. (2012). Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* 335, 233–235.
- Bartlett, J.G., Chang, T.W., Gurwith, M., Gorbach, S.L., and Onderdonk, A.B. (1978). Antibiotic-Associated Pseudomembranous Colitis Due to Toxin-Producing *Clostridia*. *N. Engl. J. Med.* 298, 531–534.
- Belogurov, G.A., and Artsimovitch, I. (2015). Regulation of Transcript Elongation. *Annu. Rev. Microbiol.* 69, 49–69.
- Berens, C., Groher, F., and Suess, B. (2015). RNA aptamers as genetic control devices: The potential of riboswitches as synthetic elements for regulating gene expression. *Biotechnol. J.* 10, 246–257.
- Blixt, T., Gradel, K.O., Homann, C., Seidelin, J.B., Schønning, K., Lester, A., Houlind, J., Stangerup, M., Gottlieb, M., and Knudsen, J.D. (2017). Asymptomatic carriers contribute to nosocomial *Clostridium difficile* infection: a cohort study of 4508 patients. *Gastroenterology*.
- Blount, K.F., Megyola, C., Plummer, M., Osterman, D., O'Connell, T., Aristoff, P., Quinn, C., Chrusciel, R.A., Poel, T.J., Schostarez, H.J., et al. (2015). Novel riboswitch-binding flavin analog that protects mice against *Clostridium difficile* infection without inhibiting cecal flora. *Antimicrob. Agents Chemother.* 59, 5736–5746.
- Bocobza, S., Adato, A., Mandel, T., Shapira, M., Nudler, E., and Aharoni, A. (2007). Riboswitch-dependent gene regulation and its evolution in the plant kingdom. *Genes Dev.* 21, 2874–2879.
- Bouillaut, L., Self, W.T., and Sonenshein, A.L. (2013). Proline-dependent regulation of *Clostridium difficile* Stickland metabolism. *J. Bacteriol.* 195, 844–854.

- Butler, E.B., Xiong, Y., Wang, J., and Strobel, S.A. (2011). Structural basis of cooperative ligand binding by the glycine riboswitch. *Chem. Biol.* 18, 293–298.
- Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the Bacterial RNA Polymerase Promoter Specificity  $\sigma$  Subunit. *Mol. Cell* 9, 527–539.
- Carey, M.F., Peterson, C.L., and Smale, S.T. (2013). The Primer Extension Assay. *Cold Spring Harb. Protoc.* 2013, pdb.prot071902-prot071902.
- Caron, M.-P., Bastet, L., Lussier, A., Simoneau-Roy, M., Massé, E., and Lafontaine, D.A. (2012). Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc. Natl. Acad. Sci. U. S. A.* 109, E3444–53.
- Caspersson, T., and Schultz, J. (1939). Pentose nucleotides in the cytoplasm of growing tissues. *Nature* 143, 602–603.
- Chanishvili, N. (2012). Phage Therapy—History from Twort and d’Herelle Through Soviet Experience to Current Approaches. In *Advances in Virus Research*, pp. 3–40.
- Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R.R. (2007). Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* 447, 497–500.
- Chen, L., Todd, R., Kiehlbauch, J., Walters, M., and Kallen, A. (2017). *Notes from the Field: Pan-Resistant New Delhi Metallo-Beta-Lactamase-Producing Klebsiella pneumoniae* — Washoe County, Nevada, 2016. *MMWR. Morb. Mortal. Wkly. Rep.* 66, 33.
- Chyung, J.W., and Shin, D.G. (2013). Localized pseudomembranous colitis in the cecum and ascending colon mimicking acute appendicitis. *World J. Gastrointest. Surg.* 5, 156–160.
- Collins, J.A., Irnov, I., Baker, S., and Winkler, W.C. (2007). Mechanism of mRNA destabilization by the glmS ribozyme. *Genes Dev.* 21, 3356–3368.
- Cornely, O.A., Miller, M.A., Louie, T.J., Crook, D.W., and Gorbach, S.L. (2012). Treatment of first recurrence of *Clostridium difficile* infection: fidaxomicin versus vancomycin. *Clin. Infect. Dis.* 55 Suppl 2, S154–61.
- Costello, S.P., Tucker, E.C., La Brooy, J., Schoeman, M.N., and Andrews, J.M. (2016). Establishing a Fecal Microbiota Transplant Service for the Treatment of *Clostridium difficile* Infection. *Clin. Infect. Dis.* 62, 908–914.
- Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561–563.
- Dahm, R. (2008). Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum. Genet.* 122, 565–581.
- Dambach, M., Sandoval, M., Updegrove, T.B., Anantharaman, V., Aravind, L., Waters, L.S., and Storz, G. (2015). The Ubiquitous yybP-ykoY Riboswitch Is a Manganese-Responsive Regulatory Element. *Mol. Cell* 57.
- Didelot, X., Eyre, D.W., Cule, M., Ip, C.L., Ansari, M., Griffiths, D., Vaughan, A., O’Connor, L., Golubchik, T., Batty, E.M., et al. (2012). Microevolutionary analysis of *Clostridium difficile* genomes to investigate transmission. *Genome Biol.* 13, R118.
- Eaton, M.D., and Bayne-Jones, S. (1934). Bacteriophage therapy: Review of the principles and results of the uses of bacteriophage in the treatment of infections. *J. Am. Med. Assoc.* 103, 1769.
- Edmond, M.B. (2016). The power of poop: Fecal microbiota transplantation for *Clostridium difficile* infection. *Trans. Am. Clin. Climatol. Assoc.* 127, 71–80.

- Edwards, A.L., Reyes, F.E., Héroux, A., and Batey, R.T. (2010). Structural basis for recognition of S-adenosylhomocysteine by riboswitches. *RNA* 16, 2144–2155.
- Eiseman, B., Silen, W., Bascom, G.S., and Kauvar, A.J. (1958). Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. *Surgery* 44, 854–859.
- Ellington, A.D., and Szostak, J.W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* 346, 818–822.
- Fagan, R.P., and Fairweather, N.F. (2011). *Clostridium difficile* has two parallel and essential Sec secretion systems. *J. Biol. Chem.* 286, 27483–27493.
- Figuerola, I., Johnson, S., Sambol, S.P., Goldstein, E.J.C., Citron, D.M., and Gerding, D.N. (2012). Relapse versus reinfection: recurrent *Clostridium difficile* infection following treatment with fidaxomicin or vancomycin. *Clin. Infect. Dis.* 55 Suppl 2, S104-9.
- Fonknechten, N., Chaussonnerie, S., Tricot, S., Lajus, A., Andreesen, J.R., Perchat, N., Pelletier, E., Gouyvenoux, M., Barbe, V., Salanoubat, M., et al. (2010). *Clostridium sticklandii*, a specialist in amino acid degradation: revisiting its metabolism through its genome sequence. *BMC Genomics* 11, 555.
- Forrest, A.R.R., and Carninci, P. (2009). Whole genome transcriptome analysis. *RNA Biol.* 6, 107–112.
- Francis, M.B., Allen, C.A., Shrestha, R., Sorg, J.A., and Gerding, D. (2013). Bile Acid Recognition by the *Clostridium difficile* Germinant Receptor, CspC, Is Important for Establishing Infection. *PLoS Pathog.* 9, e1003356.
- Furukawa, K., Ramesh, A., Zhou, Z., Weinberg, Z., Vallery, T., Winkler, W.C., and Breaker, R.R. (2015). Bacterial Riboswitches Cooperatively Bind Ni<sup>2+</sup> or Co<sup>2+</sup> Ions and Control Expression of Heavy Metal Transporters. *Mol. Cell* 57, 1088–1098.
- Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.-J., Wortman, J.R., Batzoglou, S., Lee, S.-I., Baştürkmen, M., Spevak, C.C., Clutterbuck, J., et al. (2005). Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438, 1105–1115.
- Galdys, A.L., Nelson, J.S., Shutt, K.A., Schlackman, J.L., Pakstis, D.L., Pasculle, A.W., Marsh, J.W., Harrison, L.H., and Curry, S.R. (2014). Prevalence and Duration of Asymptomatic *Clostridium difficile* Carriage among Healthy Subjects in Pittsburgh, Pennsylvania. *J. Clin. Microbiol.* 52, 2406–2409.
- Gao, A., and Serganov, A. (2014). Structural insights into recognition of c-di-AMP by the ydaO riboswitch. *Nat. Chem. Biol.* 10, 787–792.
- Gelfand, M.S., Mironov, A.A., Jomantas, J., Kozlov, Y.I., and Perumov, D.A. (1999). A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet.* 15, 439–442.
- Georg, J., and Hess, W.R. (2011). cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol. Mol. Biol. Rev.* 75, 286–300.
- George, W.L. (1980). Antimicrobial agent-associated colitis and diarrhea. *West. J. Med.* 133, 115–123.
- George, R.H., Symonds, J.M., Dimock, F., Brown, J.D., Arabi, Y., Shinagawa, N., Keighley, M.R., Alexander-Williams, J., and Burdon, D.W. (1978). Identification of *Clostridium difficile* as a cause of pseudomembranous colitis. *Br. Med. J.* 1, 695.
- Gerba, C.P., Wuollet, A.L., Raisanen, P., and Lopez, G.U. (2016). Bacterial contamination of computer touch screens.

- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345.
- Green, N.J., Grundy, F.J., and Henkin, T.M. (2010). The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett.* 584, 318–324.
- Grundy, F.J., and Henkin, T.M. (1998). The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Mol. Microbiol.* 30, 737–749.
- Guérout-Fleury, A.M., Frandsen, N., and Stragier, P. (1996). Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* 180, 57–61.
- Hall, I.C., and O'Toole, E. (1935). Intestinal Flora in New-Born Infants. *Am. J. Dis. Child.* 49, 390.
- Hall, J., and Premji, A. (2015). Essential Med Notes (Toronto, Ontario, Canada: Toronto Notes for Medical Students, Inc).
- Hartig, J.S. (2010). A group I intron riboswitch. *Chem. Biol.* 17, 920–921.
- Häuser, R., Pech, M., Kijek, J., Yamamoto, H., Titz, B., Naeve, F., Tovchigrechko, A., Yamamoto, K., Szaflarski, W., Takeuchi, N., et al. (2012). RsfA (YbeB) Proteins Are Conserved Ribosomal Silencing Factors. *PLoS Genet.* 8, e1002815.
- He, M., Sebaihia, M., Lawley, T.D., Stabler, R.A., Dawson, L.F., Martin, M.J., Holt, K.E., Seth-Smith, H.M.B., Quail, M.A., Rance, R., et al. (2010). Evolutionary dynamics of *Clostridium difficile* over short and long time scales. *Proc. Natl. Acad. Sci. U. S. A.* 107, 7527–7532.
- He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D.J., Martin, M.J., Connor, T.R., Harris, S.R., Fairley, D., Bamford, K.B., et al. (2013). Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*. *Nat. Genet.* 45, 109–113.
- Heap, J.T., Pennington, O.J., Cartman, S.T., Carter, G.P., and Minton, N.P. (2007). The ClosTron: A universal gene knock-out system for the genus *Clostridium*. *J. Microbiol. Methods* 70, 452–464.
- Heap, J.T., Pennington, O.J., Cartman, S.T., and Minton, N.P. (2009). A modular system for *Clostridium* shuttle plasmids. *J. Microbiol. Methods* 78, 79–85.
- Heap, J.T., Kuehne, S.A., Ehsaan, M., Cartman, S.T., Cooksley, C.M., Scott, J.C., and Minton, N.P. (2010). The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J. Microbiol. Methods* 80, 49–55.
- Huang, L., Serganov, A., and Patel, D.J. (2010). Structural insights into ligand recognition by a sensing domain of the cooperative glycine riboswitch. *Mol. Cell* 40, 774–786.
- Human Genome Sequencing Consortium, I. (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945.
- Johansson, J. (2009). RNA Thermosensors in Bacterial Pathogens. In *Bacterial Sensing and Signaling*, (Basel: KARGER), pp. 150–160.
- Johnson, J.E., Reyes, F.E., Polaski, J.T., Batey, R.T., and Batey, R.T. (2012). B12 cofactors directly stabilize an mRNA regulatory switch. *Nature* 492, 133–137.



- Jullian-Desayes, I., Landelle, C., Mallaret, M.-R., Brun-Buisson, C., and Barbut, F. (2017). *Clostridium difficile* contamination of health care workers' hands and its potential contribution to the spread of infection: Review of the literature. *Am. J. Infect. Control* 45, 51–58.
- Karlsson, S., Burman, L.G., and Akerlund, T. (1999). Suppression of toxin production in *Clostridium difficile* VPI 10463 by amino acids. *Microbiology* 145, 1683–1693.
- Kato, T., Yoshida, H., Miyata, T., Maki, Y., Wada, A., and Namba, K. (2010). Structure of the 100S Ribosome in the Hibernation Stage Revealed by Electron Cryomicroscopy. *Structure* 18, 719–724.
- Kim, J.N., Roth, A., and Breaker, R.R. (2007). Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2'-deoxyguanosine. *Proc. Natl. Acad. Sci.* 104, 16092–16097.
- Kladwang, W., Chou, F.-C., and Das, R. (2012). Automated RNA structure prediction uncovers a kink-turn linker in double glycine riboswitches. *J. Am. Chem. Soc.* 134, 1404–1407.
- Klein, D.J., and Ferré-D'Amaré, A.R. (2006). Structural Basis of glmS Ribozyme Activation by Glucosamine-6-Phosphate. *Science* (80-. ). 313, 1752–1756.
- Kobayashi, K., Takakuwa, H., Maegawa, T., Kato, H., Saikai, T., Kita, H., Yamagishi, T., Koino, Y., Karasawa, T., and Nakamura, S. (2001). Colonisation and transmission of *Clostridium difficile* in healthy individuals examined by PCR ribotyping and pulsed-field gel electrophoresis. *J. Med. Microbiol.* 50, 720–727.
- Kramer, A., Schwebke, I., and Kampf, G. (2006). How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infect. Dis.* 6, 130.
- Kroos, L., Zhang, B., Ichikawa, H., and Yu, Y.T. (1999). Control of sigma factor activity during *Bacillus subtilis* sporulation. *Mol. Microbiol.* 31, 1285–1294.
- Kuehne, S.A., and Minton, N.P. (2012). ClosTron-mediated engineering of *Clostridium*. *Bioengineered* 3, 247–254.
- Kwon, M., and Strobel, S.A. (2008). Chemical basis of glycine riboswitch cooperativity. *RNA* 14, 25–34.
- Kwon, J.H., Olsen, M.A., and Dubberke, E.R. (2015). The Morbidity, Mortality, and Costs Associated with *Clostridium difficile* Infection. *Infect. Dis. Clin. North Am.* 29, 123–134.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
- Langen, H., Gray, C., Röder, D., Juranville, J.-F., Takacs, B., and Fountoulakis, M. (1997). From genome to proteome: Protein map of *Haemophilus influenzae*. *Electrophoresis* 18, 1184–1192.
- Lawson, P.A., Citron, D.M., Tyrrell, K.L., and Finegold, S.M. (2016). Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. *Anaerobe* 40, 95–99.
- Lee, E.R., Baker, J.L., Weinberg, Z., Sudarsan, N., and Breaker, R.R. (2010). An Allosteric Self-Splicing Ribozyme Triggered by a Bacterial Second Messenger. *Science* (80-. ). 329.
- Leipe, D.D., Aravind, L., Grishin, N. V, and Koonin, E. V (2000). The bacterial replicative helicase DnaB evolved from a RecA duplication. *Genome Res.* 10, 5–16.
- Lesk, A.M. (1969). Why does DNA contain thymine and RNA uracil? *J. Theor. Biol.* 22, 537–540.

- Levene, P.A. (1919). The structure of yeast nucleic acid. *J. Biol. Chem.* *40*, 415–424.
- Lewin, R.A. (2001). More on Merde. *Perspect. Biol. Med.* *44*, 594–607.
- Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., Repoila, F., Buchrieser, C., Cossart, P., and Johansson, J. (2009). A trans-Acting Riboswitch Controls Expression of the Virulence Regulator PrfA in *Listeria monocytogenes*. *Cell* *139*, 770–779.
- Mandal, M., and Breaker, R.R. (2004). Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. Mol. Biol.* *11*, 29–35.
- Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., and Breaker, R.R. (2003). Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* *113*, 577–586.
- Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L., and Breaker, R.R. (2004). A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* (80-. ). *306*, 275–279.
- Marsh, J.W., Arora, R., Schlackman, J.L., Shutt, K.A., Curry, S.R., and Harrison, L.H. (2012). Association of relapse of *Clostridium difficile* disease with BI/NAP1/027. *J. Clin. Microbiol.* *50*, 4078–4082.
- McBee, R.H. (1960). Intestinal flora of some antarctic birds and mammals. *J. Bacteriol.* *79*, 311–312.
- Meador, E., Mayer, M.J., Steverding, D., Carding, S.R., and Narbad, A. (2013). Evaluation of bacteriophage therapy to control *Clostridium difficile* and toxin production in an in vitro human colon model system. *Anaerobe* *22*, 25–30.
- Michael, C.A., Dominey-Howes, D., and Labbate, M. (2014). The antimicrobial resistance crisis: causes, consequences, and management. *Front. Public Heal.* *2*, 145.
- Mitchell, B.G., and Gardner, A. (2012). Mortality and *Clostridium difficile* infection: a review. *Antimicrob. Resist. Infect. Control* *1*, 1–6.
- Montange, R.K., and Batey, R.T. (2006). Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* *441*, 1172–1175.
- Moono, P., Foster, N.F., Hampson, D.J., Knight, D.R., Bloomfield, L.E., and Riley, T. V. (2016). *Clostridium difficile* Infection in Production Animals and Avian Species: A Review. *Foodborne Pathog. Dis.* *13*, 647–655.
- Mulhbachter, J., Brouillette, E., Allard, M., Fortier, L.C., Malouin, F., and Lafontaine, D. a (2010). Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog.* *6*, e1000865.
- Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., and Breaker, R.R. (2002). Genetic Control by a Metabolite Binding mRNA. *Chem. Biol.* *9*, 1043–1049.
- Nale, J.Y., Spencer, J., Hargreaves, K.R., Buckley, A.M., Trzapiński, P., Douce, G.R., and Clokie, M.R.J. (2016). Bacteriophage Combinations Significantly Reduce *Clostridium difficile* Growth In Vitro and Proliferation In Vivo. *Antimicrob. Agents Chemother.* *60*, 968–981.
- Nechooshtan, G., Elgrably-Weiss, M., Sheaffer, A., Westhof, E., and Altuvia, S. (2009). A pH-responsive riboregulator. *Genes Dev.* *23*, 2650–2662.
- Nelson, J.W., Sudarsan, N., Furukawa, K., Weinberg, Z., Wang, J.X., and Breaker, R.R. (2013). Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nat. Chem. Biol.* *9*, 834–839.

- Nisman, B. (1954). The Stickland reaction. *Bacteriol. Rev.* *18*, 16–42.
- Nolan, N.P.M., Kelly, C.P., Hiumphreys, H., Cooney, C., O ’connor, R., Walsh, T.N., Weir, D.G., and O ’briain, D.S. (1987). An epidemic of pseudomembranous colitis: Importance of person to person spread. *Glut* *28*, 1467–1473.
- Nurmi, E., and Rantala, M. (1973). New aspects of *Salmonella* infection in broiler production. *Nature* *241*, 210–211.
- O’Connor, J.R., Lyras, D., Farrow, K.A., Adams, V., Powell, D.R., Hinds, J., Cheung, J.K., and Rood, J.I. (2006). Construction and analysis of chromosomal *Clostridium difficile* mutants. *Mol. Microbiol.* *61*, 1335–1351.
- Oliver, S.G., Winson, M.K., Kell, D.B., and Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends Biotechnol.* *16*, 373–378.
- Ozaki, E., Kato, H., Kita, H., Karasawa, T., Maegawa, T., Koino, Y., Matsumoto, K., Takada, T., Nomoto, K., Tanaka, R., et al. (2004). *Clostridium difficile* colonization in healthy adults: transient colonization and correlation with enterococcal colonization. *J. Med. Microbiol.* *53*, 167–172.
- Pedrolli, D., Langer, S., Hobl, B., Schwarz, J., Hashimoto, M., and Mack, M. (2015). The *ribB* FMN riboswitch from *Escherichia coli* operates at the transcriptional and translational level and regulates riboflavin biosynthesis. *FEBS J.* *282*, 3230–3242.
- Price, I.R., Gaballa, A., Ding, F., Helmann, J.D., and Ke, A. (2015). Mn<sup>2+</sup>-Sensing Mechanisms of yybP-ykoY Orphan Riboswitches. *Mol. Cell* *57*.
- Ramesh, A., and Winkler, W.C. (2010). Magnesium-sensing riboswitches in bacteria. *RNA Biol.* *7*, 77–83.
- Ramirez, J. (2016). Exploration of Putative Riboswitches Involved in Bacteriophage Temperate Life Cycle (aaas).
- Rashid, T., Vonville, H., Hasan, I., and Garey, K.W. (2017). Mechanisms for floor surfaces or environmental ground contamination to cause human infection: a systematic review. *Epidemiol. Infect.* *145*, 347–357.
- Ray-Soni, A., Bellecourt, M.J., and Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annu. Rev. Biochem.* *85*, 319–347.
- Regulski, E.E., Moy, R.H., Weinberg, Z., Barrick, J.E., Yao, Z., Ruzzo, W.L., and Breaker, R.R. (2008). A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Mol. Microbiol.* *68*, 918–932.
- Ren, A., Xue, Y., Peselis, A., Serganov, A., Al-Hashimi, H.M., and Patel, D.J. (2015). Structural and Dynamic Basis for Low-Affinity, High-Selectivity Binding of L-Glutamine by the Glutamine Riboswitch. *Cell Rep.* *13*, 1800–1813.
- Robertson, M.P., and Joyce, G.F. (2012). The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* *4*.
- Roßmanith, J., and Narberhaus, F. (2016). Exploring the modular nature of riboswitches and RNA thermometers. *Nucleic Acids Res.* *44*, 5410–5423.
- Roth, A., and Breaker, R.R. (2009). The Structural and Functional Diversity of Metabolite-Binding Riboswitches. *Annu. Rev. Biochem.* *78*, 305–334.
- Roth, A., Winkler, W.C., Regulski, E.E., Lee, B.W.K., Lim, J., Jona, I., Barrick, J.E., Ritwik, A., Kim, J.N., Welz, R., et al. (2007). A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.* *14*, 308–317.

- Ruff, K.M., and Strobel, S.A. (2014). Ligand binding by the tandem glycine riboswitch depends on aptamer dimerization but not double ligand occupancy. *RNA* 20, 1775–1788.
- Ruff, K.M., Muhammad, A., Mccown, P.J., Breaker, R.R., and Strobel, S.A. (2016). Singlet glycine riboswitches bind ligand as well as tandem riboswitches. *RNA* 22, 1–11.
- Sabuncu, E., David, J., Bernède-Bauduin, C., Pépin, S., Leroy, M., Boëlle, P.-Y., Watier, L., Guillemot, D., Schuchat, A., Robinson, K., et al. (2009). Significant Reduction of Antibiotic Use in the Community after a Nationwide Campaign in France, 2002–2007. *PLoS Med.* 6, e1000084.
- Sayamov, R.M. (1963). Treatment and prophylaxis of cholera with bacteriophage. *Bull. World Health Organ.* 28, 361–367.
- Schless, R.A., Griffith, J. P. C., and Mitchell, A.G., Porter, Langley, and Carter, W., Holt, L.E., Emerson, K., Ballenger, H.C., Garber, Frank W., J., Sciaroni, G.H., Wharry, H.M., Salvani, C., et al. (1932). *Staphylococcus aureus* meningitis treatment with specific bacteriophage. *Am. J. Dis. Child.* 44, 813.
- Sebahia, M., Wren, B.W., Mullany, P., Fairweather, N.F., Minton, N., Stabler, R., Thomson, N.R., Roberts, A.P., Cerdeño-Tárraga, A.M., Wang, H., et al. (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat. Genet.* 38, 779–786.
- Serganov, A., and Nudler, E. (2013). A decade of riboswitches. *Cell* 152, 17–24.
- Serganov, A., and Patel, D.J. (2009). Amino acid recognition and gene regulation by riboswitches. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1789, 592–611.
- Serganov, A., Huang, L., and Patel, D.J. (2009). Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature* 458, 233–237.
- Sherman, E.M., Esquiaqui, J., Elsayed, G., and Ye, J.-D. (2012). An energetically beneficial leader-linker interaction abolishes ligand-binding cooperativity in glycine riboswitches. *RNA* 18, 496–507.
- Shine, J., and Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. U. S. A.* 71, 1342–1346.
- Singh, S., Jing, E., and Stollman, N. (2016). Self-Limited Sepsis Syndrome Following Fecal Microbiota Therapy for Refractory *C. difficile* Infection. *Dig. Dis. Sci.* 61, 2488–2491.
- Smith, K.D., Lipchick, S. V, Ames, T.D., Wang, J., Breaker, R.R., and Strobel, S.A. (2009). Structural basis of ligand binding by a c-di-GMP riboswitch. *Nat. Struct. Mol. Biol.* 16, 1218–1223.
- Soutourina, O.A., Monot, M., Boudry, P., Saujet, L., Pichon, C., Sismeiro, O., Semenova, E., Severinov, K., Le Bouguenec, C., Coppée, J.-Y., et al. (2013). Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. *PLoS Genet.* 9, e1003493.
- Spigaglia, P. (2016). Recent advances in the understanding of antibiotic resistance in *Clostridium difficile* infection. *Ther. Adv. Infect. Dis.* 3, 23–42.
- Ster, C., Allard, M., Boulanger, S., Lamontagne Boulet, M., Mulhbach, J., Lafontaine, D., Marsault, É., Lacasse, P., and Malouin, F. (2013). Experimental treatment of *Staphylococcus aureus* bovine intramammary infection using a guanine riboswitch ligand analog. *J. Dairy Sci.* 96, 1000–1008.

- Sudarsan, N., Lee, E.R., Weinberg, Z., Moy, R.H., Kim, J.N., Link, K.H., and Breaker, R.R. (2008). Riboswitches in Eubacteria Sense the Second Messenger Cyclic Di-GMP. *Science* (80-. ). 321.
- Taur, Y., and Pamer, E.G. (2014). Fixing the microbiota to treat *Clostridium difficile* infections. *Nat. Publ. Gr.* 20, 246–247.
- Tezuka, T., and Ohnishi, Y. (2014). Two glycine riboswitches activate the glycine cleavage system essential for glycine detoxification in *Streptomyces griseus*. *J. Bacteriol.* 196, 1369–1376.
- Twort, F. (1915). An investigation on the nature of ultramicroscopic viruses. *Lancet* 11, 1241.
- Vaishnavi, C. (2015). Fidaxomicin - the new drug for *Clostridium difficile* infection. *Indian J. Med. Res.* 141, 398.
- Van Valen, L. (1973). A new evolutionary law. *Evol. Theory* 1, 1–30.
- Vassilyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassilyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719.
- Ventola, C.L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *P T* 40, 277–283.
- Vojcic, L., Despotovic, D., Martinez, R., Maurer, K.-H., and Schwaneberg, U. (2012). An efficient transformation method for *Bacillus subtilis* DB104. *Appl. Microbiol. Biotechnol.* 94, 487–493.
- Voth, D.E., and Ballard, J.D. (2005). *Clostridium difficile* toxins: mechanism of action and role in disease. *Clin. Microbiol. Rev.* 18, 247–263.
- Wachsmuth, M., Findeiss, S., Weissheimer, N., Stadler, P.F., and Morl, M. (2013). De novo design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Res.* 41, 2541–2551.
- Wachter, A., Tunc-Ozdemir, M., Grove, B.C., Green, P.J., Shintani, D.K., and Breaker, R.R. (2007). Riboswitch Control of Gene Expression in Plants by Splicing and Alternative 3' End Processing of mRNAs. *Plant Cell* 19, 3437–3450.
- Wacker, A., Buck, J., Mathieu, D., Richter, C., Wö Hnert, J., and Schwalbe, H. (2011). Structure and dynamics of the deoxyguanosine- sensing riboswitch studied by NMR-spectroscopy. *Nucleic Acids Res.* 39, 6802–6812.
- Wang, J.X., Lee, E.R., Morales, D.R., Lim, J., and Breaker, R.R. (2008). Riboswitches that Sense S-adenosylhomocysteine and Activate Genes Involved in Coenzyme Recycling. *Mol. Cell* 29, 691–702.
- Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 737–738.
- Watt, M., Dinh, A., Le Monnier, A., and Tilleul, P. (2017). Cost-effectiveness analysis on the use of fidaxomicin and vancomycin to treat *Clostridium difficile* infection in France. *J. Med. Econ.* 1–9.
- Weaver, W. (1970). Molecular Biology: Origin of the Term. *Science* (80-. ). 170.
- Weinberg, Z., Wang, J.X., Bogue, J., Yang, J., Corbino, K., Moy, R.H., Breaker, R.R., Roth, A., Breaker, R., Waters, L., et al. (2010). Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* 11, R31.

- Welz, R., and Breaker, R.R. (2007). Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*. *RNA* *13*, 573–582.
- Wenk, H.R., Downing, K.H., Hu, M., and O’Keefe, M.A. (1992). 3D structure determination from electron-microscope images: electron crystallography of staurolite. *Acta Crystallogr. A*. *48* ( Pt 5), 700–716.
- Winkler, W., Nahvi, A., and Breaker, R.R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* *419*, 952–956.
- Woodworth, M.H., Neish, E.M., Miller, N.S., Dhere, T., Burd, E.M., Carpentieri, C., Sitchenko, K.L., and Kraft, C.S. (2017). Laboratory Testing of Donors and Stool for Fecal Microbiota Transplantation for Recurrent *C. difficile* Infection. *J. Clin. Microbiol.* JCM.02327-16.
- Wright, G.D. (2010). Q&A: Antibiotic resistance: where does it come from and what can we do about it? *BMC Biol.* *8*, 123.
- Yutin, N., and Galperin, M.Y. (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ. Microbiol.* *15*, 2631–2641.
- Zeigler, D.R., Prágai, Z., Rodriguez, S., Chevreux, B., Muffler, A., Albert, T., Bai, R., Wyss, M., and Perkins, J.B. (2008). The origins of 168, W23, and other *Bacillus subtilis* legacy strains. *J. Bacteriol.* *190*, 6983–6995.
- Zhang, F., Luo, W., Shi, Y., Fan, Z., and Ji, G. (2012). Should We Standardize the 1,700-Year-Old Fecal Microbiota Transplantation? *Am. J. Gastroenterol.* *107*, 1755–1755.
- Zhang, Y., Mandava, C.S., Cao, W., Li, X., Zhang, D., Li, N., Zhang, Y., Zhang, X., Qin, Y., Mi, K., et al. (2015). HflX is a ribosome-splitting factor rescuing stalled ribosomes under stress conditions. *Nat. Struct. Mol. Biol.* *22*, 906.
- Ziervogel, B.K., and Roux, B. (2013). The binding of antibiotics in OmpF porin. *Structure* *21*, 76–87.